Reply to Referee 1

We appreciate the reviewer provided these important comments help us improving our manuscript. We'd like to address these comments as following.

First, as I read into the details, I saw in lines 182 through 184 that only one case was used to train the convective portion of the technique…. This seems very small… if this is true, how many cases (not just radar tilts) were used to train the stratiform portion of the SVM technique… This is one of the chief reasons I rated the scientific quality as only fair…. Otherwise, I would've rated it higher.

Response:
We appreciate the reviewer pointing this out. As we mentioned in the manuscript, convective type training data is mainly from a strong convective precipitation event on 23 July 2014. This thunderstorm is clearly identified as a convective precipitation through ground observation and MRMS results, that is the major reason we chose this case. Besides this case, some convective radar gates from the precipitation event on 30 August 2011 are also used as the convective training data. These convective radar gates are identified by the MRMS results. We clarified these in the revised manuscript. Total 17281 sets of data (including convective and stratiform types) are used in the training. We believe this is a reasonable number of training data at the stage of prototype algorithm development. High more numbers of training data and support vectors should be used if this approach is used for the operation purpose.

Also, the overall readability and the English needs to be improved significantly… The reason is there are statement written that will convey to the reader an opposite meaning than what the authors clearly intended. An example, can be found in line 227… where the authors mention that POD, FAR, CSI can not capture the performance of the SVM technique… If that were true why bother with those statistics …., However, I think the authors really meant that these statistics only partially capture the performance and therefore a new performance evaluation statistics should also be considered. The same subject was also mentioned in the conclusion in line 332 through 335 and should be clarified there as well… where it significantly caused me to wonder why the stats were even used …, Again, I know the authors did not intend that meaning…, but the way it is written conveyed that type of meaning…

Response:
Thank you. We totally agree with the reviewer that there are some inconsistent statements in the manuscript. In the revised manuscript, we double checked the statements, grammar, and other English issues, and corrected these errors.

We agree with the reviewer that those traditional statistics only partially capture the performance of the proposed approach. A new performance evaluation score is therefore used as a complement. We corrected this in the revised manuscript.

While I think the authors significantly improved the amount of detail needed for publication and that there is significant scientific merit to the work. I just do not think it's ready for publication yet without significant major revisions to improve the readability of the paper. However, I definitely think this paper should be revised because I believe it does carry information that would be useful for the scientific community once the revisions are made. While I know this may be discouraging for authors, they need to realize that important that they keep moving forward on this paper because persistence will ultimately in excellence.

Response:
We do appreciate the reviewer's comments, which helps us making this paper more solid. In the revised manuscript, we double checked the statements, grammar, and other English issues, and corrected these errors.

Reply to Referee 2

We appreciate the reviewer provided these important comments help us improving our manuscript. We'd like to address these comments as following.

1. Thanks for carrying out the sensitivity study. I have the feeling that it may better to remove this portion as it stands right now. In particular the scores for different ZDR bias indicates the best scores for the BAL-0.5 assuming a bias of ZDR=-0.2 dB version. So is this an indication, that there is actually a ZDR bias in the underlying data?

Response: Following the reviewer's suggestion, the section of sensitivity analysis was removed in the latest version.

We agree with the reviewer that $BAL^{-0.5}$ shows the best CSI and POD scores when an artificial -0.2 dB bias was added on the ZDR field. However, we do not think we can affirm that the actual ZDR is biased (-0.5 dB) for the following two reasons:

1.) With the manually added negative ZDR biases, BAL^{0} and $BAL^{-0.5}$ will classify more echoes as convective type. Based on the separation index calculation (Equations 2~ 5), it could be found that the separation index becomes larger with smaller ZDR value for a given reflectivity. As a result, radar echo associated with larger separation index is more likely be classified as convective type. Similar results also could be found in the simulation. That is the major reason that POD and CSI are significantly enhanced. However, on the other hand, we also can find that $BAL^{-0.5}$ with +0.2 dB ZDR bias produces better RCS score comparing to -0.2 dB ZDR bias. Therefore, +0.2 dB ZDR bias produce better RCS, and -0.2dB bias produce better CSI and POD. We can't say which one is better.

2.) On the other hand, ZDR is directly used as an input in the SVM approach. From the sensitivity test (Figure 12), we can find that ZDR with +0.2 dB bias produce better scores (RCS, POD, CSI, and FAR) than -0.2 dB bias.

Since the ZDR value is directly or indirectly used in both approaches, and we got two opposite trends, we can't draw a convincible conclusion of positive or negative bias. We appreciate the reviewer emphasis the concerns about the ZDR bias, and we know the ZDR bias play an important role in the radar echo classification. When we apply this approach into operation, we need to be very careful about the data quality control. We included more discussion in the conclusion part.

2. L 16: relatively low R: please provide a number to get an idea on what is considered low T'R .

Response: Two typical R-Z relations are well used the rainfall rate estimation: $Z = 200R^{1.6}$(stratiform), and $Z = 300R^{1.4}$ (convective). Given the fact that the reflectivity from a stratiform precipitation can't above 40 dBZ. Therefore, the maximum precipitation rate for stratiform precipitation is 11 m/hour.
Following the reviewer's suggestion, we added this value into the manuscript.

3. State why the QPE accuracy is better if one makes a distinction into those regimes.

Response: Different reflectivity-rain rate relations are generally used in QPE (Kirsch et al. 2019). Therefore, accurate classification help choosing the optimum R(Z) relation, which can produce accurate rainfall estimation results.
Following the reviewer's comment, we added this in the revised manuscript.

L40. "An alternative scanning scheme …." Something seems wrong in the sentences (grammar).

Response: This sentence is modified as:
"New radar scanning schemes are designed to update data from low elevations in a high frequency and data from high elevations in a low frequency. Such alternative scanning scheme enables the WSR-88D radars to promptly capture the storm development, which can enhance the weather forecast capability and QPE accuracy."

L58. The use of lowest unblocked sweep is used. Don't you have to live with the lowest sweep, even if it is partially blocked. You don't have another sweep available. See I38

Response:
If the lowest tilt data is partially/totally blocked, the reflectivity and differential reflectivity values become biases/unavailable. Therefore, we have to use the next adjacent low tilt data. We believe the sentence in L38 misleading, and we modified the sentence in the revised manuscript.

L107: Threshold of 0.9 is too high to discriminate clutter and meteorology. You will loose valid meteorological data. I think I addressed this already in my initial review. You have to show or discuss why this choice doesn't affect your results. I would expect that you will have problems with your convective data set.

L 186: really a rhohv > 0.98?

Response: These two comments are both related to the ρ_{HV} thresholds in the SVM approach training and testing. We do appreciate the reviewer bringing this back to us, since this is a very important issue in the data quality control. We tested a new set of thresholds (ρ_{HV} = 0.9 in training and ρ_{HV} = 0.85 in quality control) in the algorithm and clarify this issue in the manuscript.

Following the reviewer's suggestions, we did the following modifications in our SVM approach.

1.) In the training data selection, a threshold of 0.9 is applied to replace 0.98. This threshold included more radar echoes in the training data. All parameters and coefficients used in the SVM approach are changed according to the new training data. These new settings were used in the validation cases. Following the work reported in Park et al. (2009), the lowest ρ_{HV} from typical liquid phase precipitation (including light to heavy rains) is 0.92. Since the goal of this

work is proposing a liquid phase precipitation classification method, we believe 0.9 is a reasonable threshold in the training data selection.

2.) In the validation, a ρ_{HV} threshold of 0.85 is used to remove those echoes not associated with rainfall. Radar echoes associated with ρ_{HV} larger than 0.85 are remained in the classification. We understood that other meteorological targets such as wet snow, crystals, graupel may be associated with low ρ_{HV} (< 0.85). However, this work proposes an approach to classify precipitations into either stratiform or convective types. Classification of other meteorological targets (snow, graupel) is not in the scope of this work. As reported in Park et al. (2009), the ρ_{HV} threshold of 0.9 is good for the majority of the liquid phase precipitation. The ρ_{HV} from the mix of rain and hail is also above 0.85. In this work, any pixel associated with low ρ_{HV} (< 0.85) is set as null.

In the revised manuscript, Figures 4 ~ 10 were regenerated using the new thresholds ($\rho_{HV} = 0.98$ for training, and $\rho_{HV} = 0.9$ for quality control). Overall, the new thresholds do not significantly affect the results. A comparison of SVM results using old threshold set ($\rho_{HV} = 0.98$ for training, and $\rho_{HV} = 0.9$ for quality control), and new threshold set ($\rho_{HV} = 0.9$ for training, and $\rho_{HV} = 0.85$ for testing) is shown in Figure 1. In this example, we could find that major parts from these two threshold sets show similar classification results, and only slight differences could be found within the red circles.

Based on our analysis, similar results may come from two reasons:
First one is the overall high ρ_{HV} field. These cases used in this work are from pure liquid precipitation events. No apparent hails contamination (within convective) could be identified from these cases. The model sounding data indicates that the bright band is about 5.5 km above the surface, which can eliminate the bright band effect. Therefore, high ρ_{HV} field are generally observed from these cases. Therefore, the adjustment of the ρ_{HV} threshold does not have significant effects on the classification results. An example of the ρ_{HV} field is shown Fig. 2, where the ρ_{HV} values from most parts of the fields are above 0.85.

Second reason may exist in the data processing. In the proposed approach, the reflectivity and differential reflectivity fields are also further smoothed with a 3 by 3 windows after attenuation correction. This step can further decrease the variations from some individual pixels. Therefore, although some individual pixel associated with low ρ_{HV} is eliminate by threshold, the smoothing average still can fill in. To demonstrate this point, the reflectivity (Fig. 3), differential reflectivity (Fig. 4), and separation index (Fig. 5) fields processed using old threshold ($\rho_{HV} = 0.9$) and new threshold ($\rho_{HV} = 0.85$) are shown in panels A and B, respectively. Based on these comparisons, we can understand why the classification results from previous and new thresholds are similar.

Overall, we do appreciate the reviewer emphasis this issue again to us, which makes us go back to investigate the impacts of the ρ_{HV} thresholds. We added more discussion about the ρ_{HV} threshold selection. We also realize that different thresholds may cause other potential issues. In this work, we focus on proposing a classification approach. The statistical analysis about the performance of this approach including the sensitivity analysis will be reported in coming manuscripts.

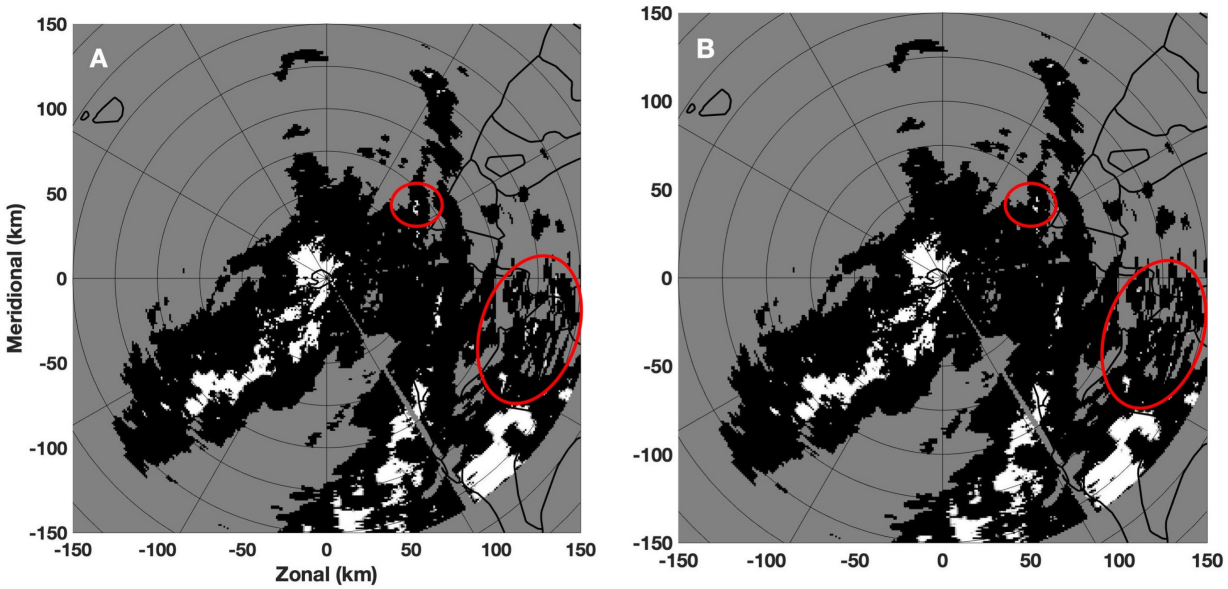

Figure 1. the SVM results using: A) previous thresholds (0.98 and 0.9) and B.) new thresholds (0.9 and 0.85).

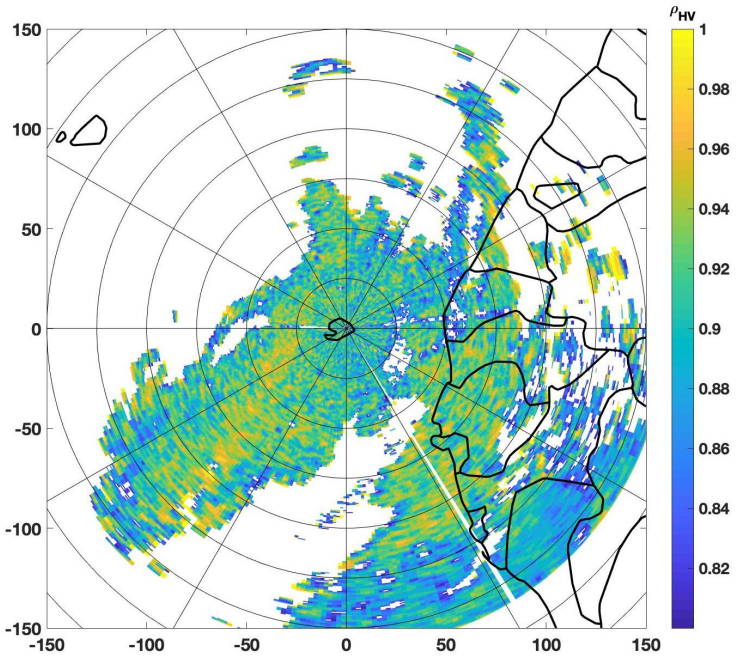

Figure 2. the ρ_{HV} field.

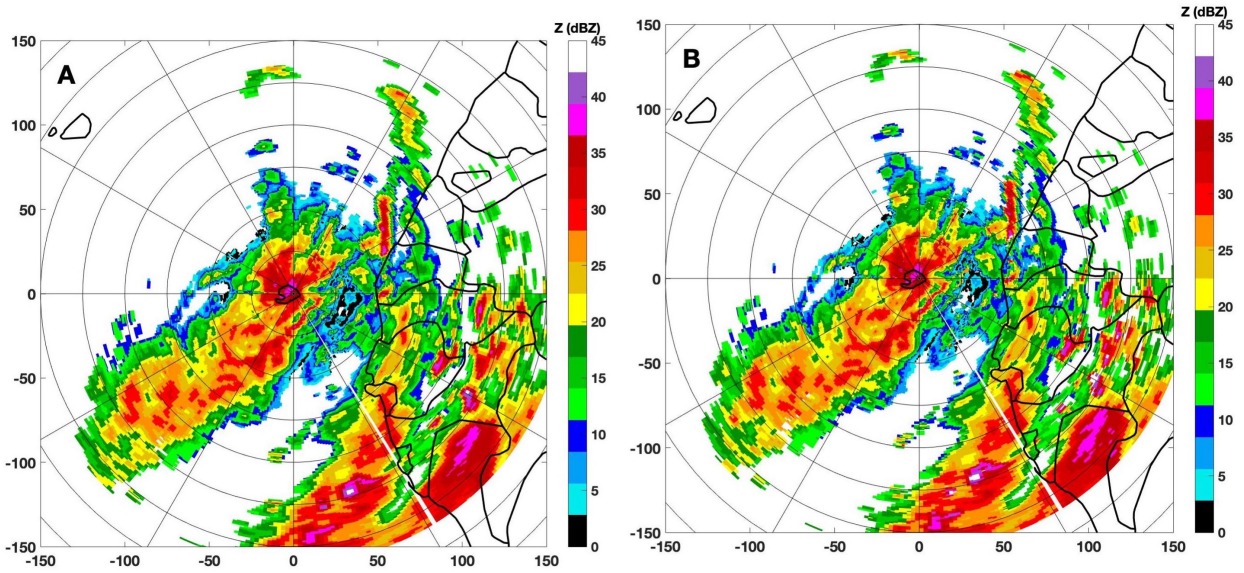

Figure 3. the smoothed reflectivity using: A) previous thresholds (0.98 and 0.9) and B.) new thresholds (0.9 and 0.85).

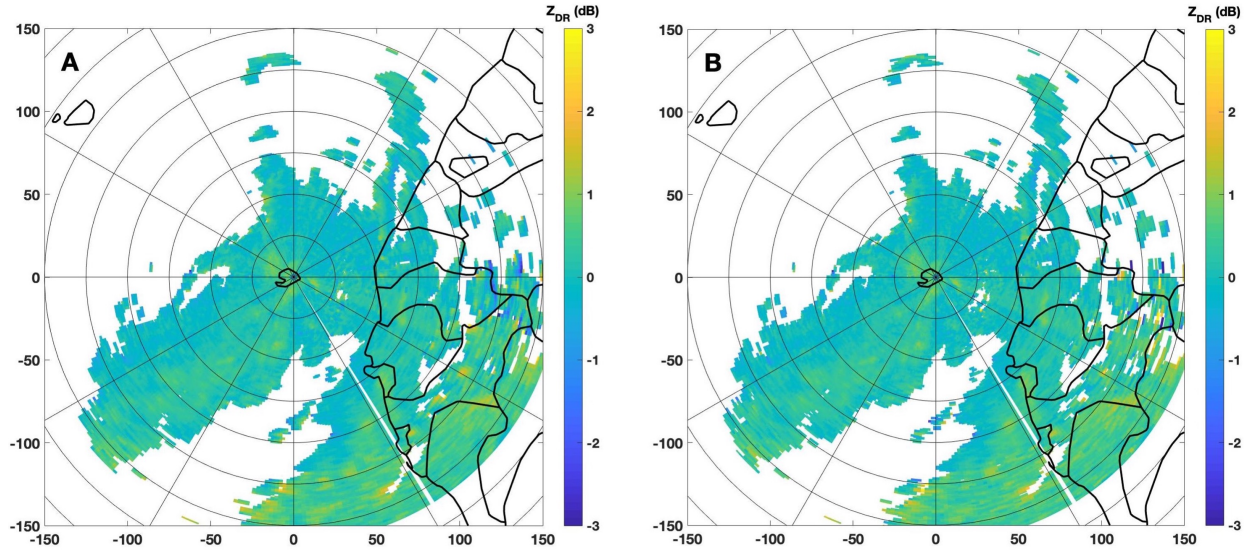

Figure 4. the smoothed differential reflectivity using: A) previous thresholds (0.98 and 0.9) and B.) new thresholds (0.9 and 0.85).

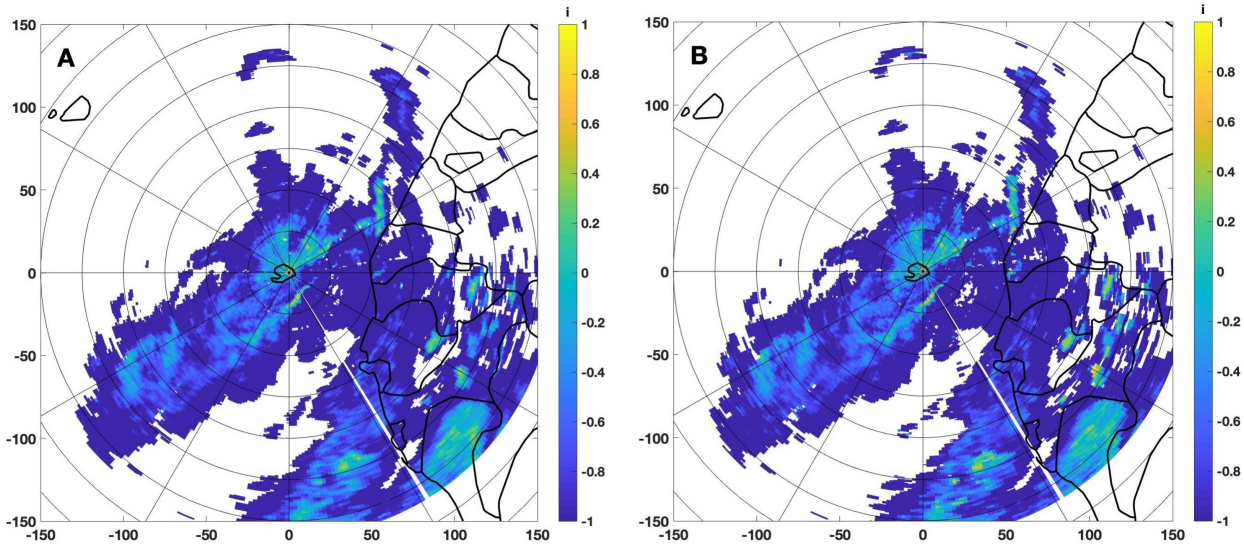

Figure 5. the separation index using: A) previous thresholds (0.98 and 0.9) and B.) new thresholds (0.9 and 0.85).

L 165: if you use a rhohv threshold of 0.9, you will exclude already a good part of the bright band data?

Response: Yes, we agree with the reviewer that ρ_{HV} threshold of 0.9 could take care of major part of the bright band. Moreover, we investigated the model sounding data for these cases. It was found that the 0^{o} temperature height is around 5 to 5.5 km in these cases. Given the fact of the low elevation angle (1.4°) and maximum range (150 km), all data is well below bright band. In this work, we only use bright band as one of the criteria to identify stratiform type precipitation, and $\rho_{HV} = 0.9$ will not applied to remove bright band signature.

L 205: Would be good to state again what those thresholds represent…. Actually, I tried to find the meaning of T_0 in the previous sections and I couldn't find it. Please clarify what does a threshold of -0.5 physically actually mean?

Response:
Using separation index only to classify precipitation into either stratiform or convective was originally proposed by Bringi et al. (2009). In their work, a threshold of 0 was suggested, and the radar echo from a gate could be classified as stratiform ($i < 0$) or convective ($i > 0$). In this work, we first tested the performance of BAL using 0 as the threshold, and it was found that a lot of radar echoes associated with convective features were classified as stratiform. Therefore, we tested a very aggressive threshold of -0.5, which classify more echoes as convective.

Following the reviewer's suggestion, the following sentence is added in the revised manuscript:

"The threshold of T_0 = 0 was suggested by BAL to classify precipitation into stratiform or convective types. In the current work, a more aggressive threshold of T_0 = -0.5 is also used in the evaluation, which will classify more radar echoes as convective than threshold of T_0 = 0."

Fig5: Why do you consider only the area in the white circle? What about the rest?

Response:
In this example, we can find that BAL^0 /$BAL^{-0.5}$ classifies the least/most echoes as convective, and SVM shows the best results comparing to BAL approach. The classification difference could be found from different locations. Here we only use the region within the circle to analyze why these three approaches show different results. Similar reasons may be applied to other regions. That is the reason that we did not discuss the region out of the circle.

L 265: you rightly state that the potential data quality problems. As such, I wonder to what extent a comparison of the methods is really meaningful. Here the SVM seems to outperform the BAL methods, correct? On the other hand, the CSI is rather poor for all methods, correct.

Response:
Thank you for reviewer pointing this out, this is a really good comment.
As we presented in the manuscript, using separation index in the precipitation classification is a good option in the fast scan fast update radar operation scheme. However, single separation index may be affected by the uncertainties in the radar field and DSD assumption. Therefore, the classification approach is potentially improved through integrating more variables. The motivation of this work is proposing a complementary method to enhance the performance of BAL, and we did identify enhancements in different precipitation events. Although scores of RCS, CSI, POD and FAR used in the manuscript show improvements in these cases, we still think these are still at the qualitative evaluation stage (Line 308 in the manuscript). These scores are rather used to demonstrate the advantages of proposed approach than statistical analysis. In order to get statistical results, we believe more long-term precipitation events are needed. The purpose of this work is introducing this SVM approach, and statistical analysis will be addressed in the future work.

We agree with the reviewer that the CSI scores are low for all methods during the comparison with MRMS. Two factors could be the major reasons: First, MRMS results are derived using the mosaicked field from four S-band radar, and the classification results are produced very 10 minutes, and other three approaches derived from RCMK, and the time gaps between MRMS and these three approaches could be as large as 5 minutes. Second, a convective storms' size, intensity, and convective cell locations could change dramatically during a short period. Therefore, pixel-to-pixel based evaluation scores may lose tracking the performance of the proposed approach, which makes the CSI showing low value. However, overall, the scores from SVM shows higher values comparing to BAL.

L 285: why don't you show the scores from the stratiform event.

Response:
This is a pure stratiform precipitation event. No convective cells are identified by all these approaches. The POD = 100%, CSI = 100%, and FAR = 0. Therefore, no scores and plots were shown in the manuscript. We added these scores in the revised manuscript.

L 295: ZDR calibration bias ….: plus the over correction due to the attenuation correction. But I don't understand the reasoning that positive zdr bias should lead to more stratiform classification.

Response:
The separation index is calculated using Z, ZDR and an DSD assumption (Equations 2~ 5). The calculated separation index shows larger value with a smaller ZDR value for a given reflectivity. The classification result is based on the comparison between calculated separation index and a threshold. A larger separation index is more likely be classified as convective type, and smaller separation index is more likely be classified as stratiform type. Therefore, when a positive bias is added into the ZDR field, the ZDR value becomes bigger, and the calculated separation index will show smaller values. As the results, more radar echoes will be classified as stratiform type.

Reference:

Kirsch, B., M. Clemens, F. Ament, 2019: Stratiform and convective radar reflectivity-rain rate relationships and their potential to improve radar rainfall estimation. *Journal of Applied Meteorology and Climatology,* vol 58, 2259-2271

LIST OF CHANGES MADE IN THE MAUSCRIPT

1.) Remove the sensitivity test section.
2.) Change the thresholds of ρ_{HV} used in the training data selection and quality control. Regenerate all results using the new thresholds.
3.) Add more discussions required by the reviewers in the revised manuscript.
4.) Modify/rewrite sentences suggested by the reviewers.

Separation of Convective and Stratiform Precipitation Using Polarimetric Radar Data with A Support Vector Machine Method

Yadong Wang[1], Lin Tang[2], Pao-Liang Chang[3], and Yu-Shuang Tang[3]

[1]Electrical and Computer Engineering Department, Southern Illinois University Edwardsville, Illinois, USA
[2]Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma, USA
[3]Central Weather Bureau, Taipei, Taiwan

Correspondence: Yadong Wang (yadwang@siue.edu)

Abstract. A precipitation separation approach using a support vector machine method was developed and tested on a C-band polarimetric weather radar located in Taiwan (RCMK). Different from some existing methods requiring a whole volume scan data, the proposed approach utilizes ~~the~~ polarimetric radar data from the lowest unblocked tilt to classify precipitation echoes into either stratiform or convective type. ~~The~~ In this algorithm, inputs of radar reflectivity, differential reflectivity, and the separation index are integrated ~~into the classification~~ through a support vector machine~~algorithm. The feature vector and weight vector~~ . The weight vector and bias in the support vector machine were optimized using well-classified training data. The proposed approach was tested with ~~multiple~~ three precipitation events including a widespread mixed stratiform and convective event, a tropical typhoon precipitation event, and a stratiform precipitation event. ~~In the evaluation, the results from the multi-radar-multi-sensor~~ Results from the Multi-Radar-Multi-Sensor (MRMS) precipitation classification algorithm were used as the ground truth ~~. The performances from~~ in the performance evaluation. The performance of the proposed approach ~~were further~~ was also compared with the approach using the separation index only. It was found that the proposed method can accurately classify the convective and stratiform ~~precipitation~~precipitations, and produce better results than the approach using the separation index only.

1 Introduction

Convective and stratiform precipitations exhibit ~~a~~ significant difference in precipitation growth mechanisms and thermodynamic structures (e.g., Houghton, 1968; Houze, 1993, 1997). Generally, a convective precipitation is associated with strong but small areal vertical air motion (> 5 m s^{-1}) (Penide et al., 2013), and delivers a high rainfall rate (R) (Anagnostou, 2004). On the other hand, a stratiform precipitation is associated with weak updrafts/downdrafts (< 3 m s^{-1}) and relatively low R ($R < 11$ mm hr^{-1}). Classifying a precipitation into either convective or stratiform type ~~not only promotes the understanding of cloud physics but also enhances the accuracy of~~ can promote our understandings in cloud physics and thermodynamics. Moreover, since different radar reflectivity (Z) and R relationships are used in the quantitative precipitation estimation (QPE), accurate classification results also can help selecting the optimal relationship and improving the QPE accuracy. For these purposes,

numerous methods using ground in situ measurements or satellite observations were developed ~~during the past four decades~~ and applied (e.g., Leary and Jr., 1979; Adler and Negri, 1988; Tokay and Short, 1996; Hong et al., 1999).

Ground-based weather radars, such as Weather Surveillance Radar, 1988, Doppler (WSR-88D), are currently used in all aspects of weather diagnosis and analysis. Precipitation classification algorithms using single- or dual-polarization radars were developed during the past three decades. For a single-polarization radar, developed algorithms mainly rely on ~~radar reflectivity~~ ~~(~~Z ~~)~~ and its derived variables (e.g., Biggerstaff and Listemaa, 2000; Anagnostou, 2004; Yang et al., 2013; Powell et al., 2016). For example, Steiner et al. (1995) (hereafter SHY95) proposed a separation approach that utilizes the texture features derived from the radar reflectivity field. In this approach, a grid point in the Z field is identified as the convective center if its value is larger than 40 dBZ, or exceeds the average intensity taken over the surrounding background by specified thresholds. Those grid points surrounding the convective centers are classified as convective area, and far regions are classified as stratiform area. Penide et al. (2013) found that SHY95 may misclassify those isolated points embedded within stratiform precipitation or associated with low cloud-top height. Powell et al. (2016) modified the SHY95′s approach, and the new approach can identify shallow convection embedded within large stratiform regions. A neural network based convective-stratiform classification algorithm was developed by Anagnostou (2004). Six variables were used in this approach as inputs including storm height, reflectivity at 2 km elevation, the vertical gradient of reflectivity, the difference in height, the standard deviation of reflectivity, and the product of reflectivity and height. Similar variables were also used in a fuzzy logic based classification approach proposed by Yang et al. (2013).

~~Although these listed classification algorithms have been developed and validated for years, a new robust algorithm is motivated for the following two reasons. The first is to utilize only the low tilt radar data for classification. According to the~~ Motivations of developing a new classification algorithm are mainly from two aspects. First, according to the U.S. Radar Operations Center (ROC), the WSR-88D radars are currently operated without updating a complete volume during each volume scan, especially during precipitation events. New radar scanning schemes are designed to ~~reorganize the updating order for a high frequency in low elevations and a less frequency for high elevations . An~~ update data from low elevations in a high frequency and data from high elevations in a low frequency. Such alternative scanning scheme enables the WSR-88D radars ~~are able~~ to promptly capture the storm development~~for weather forecast and to obtain a more accurate precipitation estimation~~, which can enhance the weather forecast capability and QPE accuracy. These new schemes include the automated volume scan evaluation and termination (AVSET), supplemental adaptive intra-volume low-level scan (SAILS), the multiple elevation scan option for SAILS, and the mid-volume rescan of low-level elevations (MRLE). ~~Under~~ With these new scanning schemes, ~~the~~ separation of stratiform/convective becomes a challenge for those algorithms requiring a full volume scan of data. Therefore, a separation algorithm using only low tilt radar data is desired. The second reason is to further explore the applications of the polarimetric variables. Polarimetric weather radars have been well applied in radar QPE, severe weather detection, hydrometeor classification, and ~~microphysical retrievals~~ cloud microphysics retrieval (Ryzhkov and Zrnic, 2019; Zhang, 2016). Extra information about hydrometeors' size, shape, ~~species,~~ and orientation could be ~~extracted~~ obtained through transmitting and receiving electromagnetic waves along the horizontal and vertical directions. Therefore, the polarimetric measurements may reveal more precipitation's microphysical and dynamic properties. Inspired by these features, a C-band polarimetric radar

precipitation separation approach was developed by Bringi et al. (2009) (hereafter BAL), which classifies the precipitation into stratiform, convective and transition regions based on retrieved drop size distribution (DSD) characteristics. However, it was found that strong stratiform echoes might have similar DSDs to weak convective echoes and lead to wrong classification results (Powell et al., 2016).

In this work, a novel precipitation separation algorithm using the separation index with other radar variables was developed and tested on a C-band polarimetric radar located in Taiwan. This approach classifies precipitations into stratiform or convective type with a support vector machine (SVM) method. Different from some existing classification techniques that require a whole volume scan of radar data, this new approach uses the lowest unblocked tilt data in the separation. If the lowest tilt is partially or completely blocked, then the next adjacent unblocked tilt is used instead. The major advantage of this method is that it can provide ~~real-time~~ classification results even ~~if~~ when the radar is operated under AVSET, SAILS, and MRLE scanning schemes~~, where the low tilts are frequently scanned and updated~~. Under these schemes, a complete volume scan is not available. This paper is organized as follows: Section 2 introduces the proposed method including radar variables and ~~processings~~data processing, the SVM method, and the training process. The performance evaluation is shown in Section 3, and the discussion and summary are given in Section 4.

2 Precipitation Separation With a Support Vector Machine Method

In the current work, the SVM precipitation separation approach was developed and validated on a C-band polarimetric radar (RCMK) located at Makung, Taiwan (Figure 1). The Weather Wing of the Chinese Air Force deployed this radar and made the data available to the Central Weather Bureau (CWB) of Taiwan since 2009. Together with three single-polarization S-band WSR-88D (RCCG, RCKT, and RCHL) and one dual-polarization S-band radar (RCWF), these five radars provide real-time ~~QPE~~ products to CWB to support missions of flood monitoring and prediction, landslide forecasts and water resource management. ~~Operating with a wavelength of~~ The wavelength of RCMK is 5.291 cm, and its range and angular resolutions are 500 m and 1°, respectively. RCMK performs volume scans of 10 tilts (0.5°, 1.4°, 2.4°, 3.4°, 4.3°, 6.0°, 9.9°, 14.6°, 19.5°, and 25°) in every 5 minutes~~with the range resolution of 500 m and angular sampling of 1°~~.

The Central Mountain Range (CMR) of Taiwan is also shown in Figure1, which poses a major challenge for radar based products. Radars located in complex terrain are prone to partial or total blockages, which cause data from the low elevation angles (EA) to be unavailable or problematic. Blockage maps of RCMK are illustrated in Figure 2. Since there are severe blockages at the 0.5° for RCMK, data from the 1.4° EA is used in the algorithm development.

2.1 Input polarimetric radar variables and preprocesses

Three measured or derived radar variables are proposed as inputs to the SVM approach: Z, differential reflectivity fields (Z_{DR}), and separation index (i). In most of precipitation classification approaches, Z is used as one of the inputs because reflectivity ~~from convective~~ fields from convective precipitations generally show higher values than ~~from stratiform type~~the fields from

stratiform precipitations. For example, a radar echo, with the reflectivity of 40 dBZ and above, is automatically classified as convective type in the approach developed by SHY95.

Differential reflectivity, which is highly related to raindrop's mass weighted mean diameter (D_m), is another good indicator of precipitation type. It was found the values of D_m in stratiform and convective precipitation generally are within 1-1.9 mm and above 1.9 mm, respectively (Chang et al., 2009). Higher Z_{DR} values are expected from convective than from stratiform ~~precipitation~~precipitations. Therefore, the Z_{DR} field is used as another input ~~of~~ in the proposed approach.

For short wavelength radars such as C-band or X-band radars, the Z and Z_{DR} fields ~~may~~ will be significantly attenuated when radar beam propagates through heavy precipitation regions. Both Z and Z_{DR} fields need to be corrected from attenuation before applied in the precipitation classification and QPE. Different attenuation correction methods were proposed using the differential phase (ϕ_{DP}) measurement such as the linear ϕ_{DP} approach, the standard ZPHI method, and the iterative ZPHI method (e.g., Jameson, 1992; Carey et al., 2000; Testud et al., 2000; Park et al., 2005). Because of its simplicity and easy implementation in a real-time system, the linear ϕ_{DP} method was applied in the current work.

$$Z(r) = Z'(r) + \alpha(\phi_{DP}(r) - \phi_{DP}(0)) \tag{1a}$$

$$Z_{DR}(r) = Z'_{DR}(r) + \beta(\phi_{DP}(r) - \phi_{DP}(0)) \tag{1b}$$

where $Z'(r)$ ($Z'_{DR}(r)$) is the observed reflectivity (differential reflectivity) at range r; $Z(r)$ ($Z_{DR}(r)$) is the corrected value; $\phi_{DP}(0)$ is the system value; $\phi_{DP}(r)$ is the smoothed (by FIR filter) differential phase at range r. The attenuation correction coefficients α and β depend on DSD, drop size shape relations (DSR), and temperature. The typical range of α (β) is found 0.06$\sim$0.15 (0.01$\sim$0.03) dB deg^{-1} for C-band radars (e.g., Carey et al., 2000; Vulpiani et al., 2012). Following the work from Wang et al. (2014), optimal coefficients α and β in Taiwan are 0.088 dB deg^{-1} and 0.02 dB deg^{-1}, respectively. The Z and Z_{DR} fields are further smoothed with a 3 (~~azimuthal~~along azimuthal angle) by 3 (along range) moving window function after corrected from attenuation.

Other quality control issues, including calibration, reflectivity vertical profile, and ground clutter removal, were also considered in this work. Since this radar is used in the real-time ~~quantitative precipitation estimation~~QPE, the biases of Z and Z_{DR} should be within 1 dBZ, and 0.1 dB, respectively. The data quality of RCMK was examined through validating the QPE performance in different works (e.g., Wang et al., 2013, 2014). Therefore, the calibration bias of RCMK should be within the reasonable range. A vertical profile of reflectivity (VPR) correction is generally needed on the reflectivity field to reduce the measurement biases because of the melting layer (Zhang et al., 2011). Given the fact that data from 1.4° elevation angle is used within the maximum range of 150 km, and the melting layer is usually around 5 km in Taiwan, the radar data used in this work is well below the melting layer. In addition, ~~considering~~ since the vertical profile of ~~differential reflectivity~~ Z_{DR} is not well studied in the current stage, no vertical corrections are applied to fields of Z and Z_{DR}. Ground clutter is typically associated with a low correlation coefficient (ρ_{HV})~~, the~~. The ρ_{HV} threshold used in this work is ~~0.9~~0.85, which can effectively remove ~~those non-meteorological echoes~~ radar echoes from non-meteorological objects such as ground clutter (Park et al., 2009).

Using the separation index i to identify convective from stratiform precipitation was initially proposed by BAL, where i was calculated under a normalized gamma DSD assumption:

$$i = log_{10}(N_W^{est}) - log_{10}(N_W^{sep}) \tag{2}$$

$$log_{10}(N_W^{sep}) = -1.6D_0 + 6.3 \tag{3}$$

where N_W^{est} is the estimated N_W (normalized number concentration) from observed Z and Z_{DR}, and is calculated as:

$$N_W^{est} = Z/0.056D_0^{7.319} \tag{4}$$

In Equation 4, D_0 is the median volume diameter and can be calculated as.

$$D_0 = 0.0203Z_{DR}^4 - 0.1488Z_{DR}^3 + 0.2209Z_{DR}^2 + 0.5571Z_{DR} + 0.801; \quad -0.5 \leq Z_{DR} < 1.25 \tag{5a}$$

$$= -0.0355Z_{DR}^3 - 0.3021Z_{DR}^2 + 1.0556Z_{DR} + 0.6844; \quad 1.25 \leq Z_{DR} < 5 \tag{5b}$$

The units of Z_{DR}, Z, N_w, and D_0 are dB, mm^6m^{-3}, $mm^{-1}m^{-3}$, and mm, respectively. The positive and negative values of index i indicate convective and stratiform rain, respectively, and $|i| < 0.1$ indicates transition regions (Penide et al., 2013). BAL pointed out that index i worked well in most of the cases in their study; however, incorrect classification results are likely obtained for low Z and high Z_{DR} cases in some convective precipitations.

2.2 Drop size distribution and drop shape relation

It should be noted that the relations between i, N_w, and D_0 were derived using the DSD data collected in Darwin, Australia. Coefficients in Equations 2∼5 need be adjusted according to the radar frequency or/and DSD and DSR features from the specific location (Thompson et al., 2015). In the current work, the separation index i is directly derived using Equations 2∼5 without further adjustment. It was shown by Wang et al. (2013) that DSD and DSR features in Taiwan are very similar to those measured from Darwin, Australia. Similar $R(K_{DP})$ relationships were obtained using data collected from these two locations. Coefficients derived by BAL could be directly used in Taiwan without further modification. To verify this assumption, N_w and D_0 were calculated using DSD data collected ~~by~~ from four impact-type Joss-Waldvogel disdrometers (JWD) located in Taiwan (Figure 1). The measurement range and temporal resolution of these JWDs are 0.359 mm ∼ 5.373 mm and 1 minute, respectively. A total of 4306-minute data from 2011∼2014 are used in N_w and D_0 calculation following the approach described in Bringi et al. (2003). Similar to the work presented in BAL, the distribution of i along median volume diameter D_0 is shown in Figure 2, where ($log_{10}N_w$, D_0) pairs from stratiform and convective types are represented with gray circles and black stars, respectively. Although the relation described in Equation 3 can separate most stratiform from convective type, a large number of points are still classified incorrectly. Therefore, the single separation index is not sufficient in the precipitation separation, and other variables such as Z and Z_{DR} may be used as supplements.

2.3 Support vector machines (SVM) method

2.3.1 Introduction of SVM

Machine learning algorithms ~~based on~~ using meteorological radar data were well developed during the past two decades (e.g., Capozzi et al., 2018; T. et al., 2019; Yen et al., 2019). ~~Support~~ A support vector machine (SVM) can be viewed as a kernel-based machine learning approach, which nonlinearly maps the data from the low-dimension input space to a high-dimension feature space, and then linearly maps to a binary output space (Burges, 1998). Given a set of training samples, the SVM constructs an optimal hyperplane, which maximizes the margin of separation between positive and negative examples (Haykin, 2011). Specifically, given a set of training data $\{(X_i, y_i)\}_{i=1}^{N}$, the goal is to find the optimal weights vector W and a bias b such that

$$y_i(W^T X_i + b) \geq 1 \qquad i = 1, 2, ..., N \tag{6}$$

where $X_i \in \mathbb{R}^m$ is the input vector, m is the variable dimension (m = 3 in this work), N is the number of training samples, and y_i is the output with the value of $+1$ or -1 that represents convective or stratiform, respectively. The particular data points (X_i, y_i) are called support vector when Equation 6 is satisfied with the equality sign. The optimum weights vector W and bias b can be obtained through solving the Lagrangian function with the minimum cost function (Haykin, 2011).

Since the SVM can be viewed as a kernel machine, finding the optimal weight vector and bias in Equation 6 can be alternatively solved through the recursive least square estimations of:

$$\sum_{i=1}^{N_s} \alpha_i y_i k(X, X_i) = 0 \tag{7}$$

where N_s is the number of support vectors, α_i is the Lagrange multipliers, and $k(X, X_i)$ is the Mercer kernel defined as:

$$k(X, X_i) = \Phi^T(X_i)\Phi(X) = exp\left(-\frac{1}{2\sigma^2}||X - X_i||^2\right) \tag{8}$$

With the solved $\{\alpha_i\}_{i=1}^{N}$, the SVM calculate the classification results with new input data $Z \in \mathbb{R}^m$ as:

$$f(Z) = sign[\left[\sum_{i=1}^{N_s} \alpha_i y_i \Phi^T(X_i)\Phi(Z)\right] \tag{9}$$

When $f(Z) = 1$, the output is classified as convective, otherwise is classified as stratiform.

2.3.2 Training of the SVM

In the SVM approach, the weight vector and bias in Equation 6 need to be optimized through a recursive least square estimation using training data. Since the training data play a critical role in the SVM approach, Z, Z_{DR} and i from convective and stratiform precipitation events were carefully examined through three steps. Firstly, the training data was checked following general classification principles. For example, training data from convective precipitation is generally associated with relatively strong reflectivity and high vertically integrated liquid (VIL). On the other hand, stratiform precipitations are generally associated

with a prominent bright band signature. The melting hydrometeors within the bright band increase backscatter during stratiform rainfall, which can significantly enhance radar reflectivity. ~~-~~The bright band feature is one of the obvious indicators of stratiform precipitation. Bright band signature normally can be observed from relatively high EAs (such as above 9.9°). ~~From low EAs, because of the combination of radar beam broadening and low slant angle, the bright band feature spreads into more gates and becomes not apparent. Therefore, in~~ In this work, the bright band feature from high elevation angles is only used in training data selection but not used as one of the inputs. Secondly, the precipitation type is verified by ground observation, such as ground severe storm reports. Thirdly, the precipitation type is confirmed by the Multi-Radar-Multi-Sensor (MRMS) precipitation classification algorithm implemented in Taiwan (Zhang et al., 2011, 2016). In this MRMS classification approach, a three-dimensional radar reflectivity field ~~was~~ is mosaicked from 4 S-band single-polarization radars (Figure 1)~~,~~. The composite reflectivity (CREF) and other measurements such as temperature and moisture fields ~~were~~ are then used in the surface precipitation classification (Zhang et al., 2016). ~~Based on the classification results, MRMS chooses different $R(Z)$ relations in the rainfall rate estimation.~~ The performance of MRMS has been thoroughly evaluated for years ~~for the quantitative precipitation estimation~~in QPE, flash flood monitoring, severe weather, and aviation weather surveillance (e.g., Gourley et al., 2016; Smith et al., 2016). ~~The productsare~~ It was found that MRMS system can provide robust and accurate products, and these products were used as the benchmark and/or ground truth in many studies (e.g., Grecu et al., 2016; Skofronick-Jackson and Coauthors, 2017). ~~It should be noted that, on the other hand, the MRMS also shows limitations since it only uses single-polarization variables to determine the precipitation type.~~ At the current stage, the MRMS precipitation classification is considered as the appropriate benchmark in the training and validation of the proposed algorithm. Moreover, since the MRMS classification is a mosaicked product derived from 4 S-band radars, it can be viewed as an independent reference.

Convective type training data is mainly from a strong convective precipitation event on 23 July 2014. This thunderstorm, classified as convective precipitation by MRMS, was associated with strong updrafts/downdrafts and caused an aircraft crash on the airport of Makung at 1106 UTC. The squall line features can be clearly identified from this storm. Radar data collected from 1030 to 1130 UTC were used as the convective type training data. The training data selection follows the criteria of $Z > 20$ dBZ, and ~~correlation coefficient (ρ_{HV}) > 0.98 (Kumjian, 2013). The stratiform type data are from~~ $\rho_{HV} \geq 0.9$. Some convective type radar data from the a mixed stratiform and convective precipitation event on 30 August 2011 are also used as the convective training data. These convective radar data are identified by the MRMS results. The stratiform type data are from the precipitation event on 30 August 2011, and only those data identified as a stratiform type by MRMS are used in training. A total of 17281 ~~sets of data~~ data sets (15144 sets of stratiform, and 2137 sets of convective) are used in the training process. In this work, one data set is defined as the variables from a gate in terms of range and azimuthal angle. Be more specific, a set of training data means a vector of $[Z(a,r)\ Z_{DR}(a,r)\ i(a,r)\ d(a,r)]$, where a and r indicate azimuthal angle and range, respectively. The variable d is the ground truth (with 1 and -1 represents convective and stratiform), which is as the desired response in the training process. The number of support vectors is selected as 1000 in the current work, and the training process is considered as completed when the root-mean-square error reaches a stable value. In the SVM approach, the original three-dimension input space nonlinearly maps to a 1000-dimension feature space, and then linearly maps to a binary output space (Burges, 1998). The higher dimension feature space potentially captures more input variables features with higher computation

cost. Generally, after the number of support vector reaches some number, the enhancement in the SVM's performance approach becomes slight. There is a balance between accuracy and computation. In the current work, the number of support vectors was tested with a value of 500, 750, 1000, 2000, and 5000. ~~The testing of~~ Since 1000 support vectors can produce less than 5% error with reasonable computation time~~. As the prototype algorithm, the number of support vectors is selected as 1000~~ , they are used in the current work.

3 Performance Evaluation

3.1 Description of the experiments

The performance of the proposed approach was validated with three precipitation events from 2009 to ~~2011. These three precipitation events include~~ 2011 including one stratiform precipitation, one intense tropical precipitation, and one mixed convective and stratiform precipitation. Two experiments based on the BAL approach with different thresholds (i.e., BAL^{0} and $BAL^{-0.5}$) were also ~~carried out in the evaluation~~validated with the same events. In these two experiments, the separation index i from each radar gate ~~is~~ was first calculated using Equations 2∼5, and thresholds of T_0 = 0 and -0.5 ~~are~~ were then used to separate convective type from stratiform type. A gate is classified as convective ~~if~~ type if obtained i is larger than T_0, and as stratiform ~~otherwise.~~ type otherwise. It was suggested that positive (negative) i is generally associated with convective (stratiform) precipitation (Bringi et al., 2009). Therefore, T_0 = 0 was selected as one of the thresholds. Another aggressive threshold of -0.5 was also tested in the current work, which will classify much more pixels as convective. This work aims to develop a complementary method using separation index i and other variables to separate convective from stratiform type. The proposed SVM and ~~BAL methods can classify the precipitation using the lowest tilt radar data only~~the BAL methods only use the lowest tilt unblocked radar data in the classification, which is suitable for fast ~~scanning and quickly updated~~ scan and quick update purposes. Other classification approaches as introduced in section 1 were not examined in the current work, because they require the data from multiple elevation angles. Therefore, their performance is not examined in this work.

In the evaluation, three statistical scores of probability of detection (POD), false alarm rate (FAR), and critical success index (CSI) are ~~first~~ used, and MRMS classification results are used as the "ground truth" in the calculation.

$$POD = \frac{hit}{hit + miss} \tag{10}$$

$$FAR = \frac{false}{hit + false} \tag{11}$$

$$CSI = \frac{hit}{hit + false + miss} \tag{12}$$

where "hit," "false," and "miss" are defined as a radar gate classified as convective type by MRMS and the evaluated approach simultaneously, by the evaluated approach only, and by MRMS only, respectively. ~~Although these scores are well~~

used in statistical analysis, two factors make it necessary to introduce one more criterion in the evaluation. First, MRMS results are derived using the mosaicked field from four S-band single-polarization radars, and the classification results are produced every 10 minutes. On the other hand, BAL^{0} , BAL$^{-0.5}$, and SVM generate classification results whenever RCMK completes a whole scan. The time difference between results from RCMK (BAL^{0} , BAL$^{-0.5}$, and SVM) and MRMS could be as large as 5 minutes. Second, a convective storm's size, intensity, and cells locations could change significantly during a short period. Therefore, these three pixel-to-pixel based evaluation scores cannot really reflect the performance of the proposed approach. As a supplement, a Since these scores only partially capture the performance due to the time gap between MRMS and RCMK results (SVM and BAL), a new evaluation score R^{CS} (whole coverage convective ratio(R^{CS}) is introduced in the current work) is also used as a supplement:

$$R^{CS} = \frac{N^{con}}{N^{con} + N^{str}} \times 100\% \tag{13}$$

Where N^{con} and N^{str} are the total pixel numbers of convective and stratiform types within the coverage, respectively. Together with CSI, POD and FAR, these four scores are used in the performance qualitative validation. The evaluation results are shown in the following sections.

3.2 Experiment results

3.2.1 Widespread mixed stratiform and convective precipitations

The performance of the proposed approach was first validated with one widespread stratiform and convective mixed precipitation event on 30 August 2011, and 24-hour data (0000 UTC∼2400 UTC) were used in the evaluation. Classification results from the proposed SVM were calculated with the trained weight vector and biases, and results from the BAL approach (BAL^{0} and BAL$^{-0.5}$) were also calculated for the comparison purpose. It should be noted that the threshold of -0.5 is much lower than the value suggested by BAL, and BAL$^{-0.5}$ may will classify more precipitations as convective type.

The time series of R^{CS} (A), CSI (B), FOD (C), and FAR (D) are calculated using Equations 10∼13 and shown in Figure 4, where results from MRMS, SVM, BAL^{0}, and BAL$^{-0.5}$ are presented by black, red, blue, and green lines, respectively. When the MRMS results are applied as the ground truth, BAL^{0} obviously classifies more precipitation as stratiform type during this 24-hour period. The time series of R^{CS} from BAL^{0} are much lower than other three approaches. BAL$^{-0.5}$ classifies more pixels as convective than BAL^{0} as expected, and the R^{CS} scores are much higher than BAL^{0}. The proposed SVM shows the most similar R^{CS} scores to MRMS comparing to BAL approaches. Since the BAL$^{-0.5}$ uses a very low threshold, it classifies more pixels as convective type, and the obtained R^{CS} scores are higher than MRMS. In term of CSI, POD, and FAR, SVM and BAL$^{-0.5}$ show similar results, but BAL^{0} show apparently worse performance.

To better understand the performance of each approach, the classification results and radar variables (Z, Z_{DR}, and i) from two distinct moments are examined and shown in Figures 5∼7. Figure 5 shows the classification results from 0303 UTC 30 August 2011, where BAL^{0}, BAL$^{-0.5}$, SVM and MRMS are shown in the panel 'A', 'B', 'C', and 'D', respectively. The time stamp for the MRMS result is 0300 UTC, and about 3 minutes earlier than the other three approaches. These three input

variables of SVM at 0303 UTC are shown in Figure 6, where Z, Z_{DR}, and i are presented in panel 'A', 'B', and 'C'. From Figures 4 and 5, it could be found that the R^{CS} from MRMS, SVM, and BAL$^{-0.5}$ show similar values, but R^{CS} from BAL^{0} is distinctly low. Within the red circle of Figure 6, the averages of Z and Z_{DR} both show relatively large values ($Z >$ 36 dBZ and $Z_{DR} >$ 0.75 dB), this is a clear indication of convective type precipitation. Both SVM and BAL$^{-0.5}$ classify most of the area within the red circle as convective, and this result is consistent with the MRMS result. Since the separation indexes within the black circle are below or slightly higher than 0, most of the area is classified as stratiform type by BAL^{0}. For this moment, threshold -0.5 shows better performance than 0.

Figure 7 shows another example of classification results from 0650 UTC. At this moment, although SVM and BAL$^{-0.5}$ produce similar CSI (0.30 v.s. 0.25) and POD (0.48 v.s. 0.52), the R^{CS} from BAL$^{-0.5}$ (32%) is much higher than R^{CS} from MRMS (17%) and SVM (13%). These scores could also be found in Figure 4. In Figure 7, It could be found from that the MRMS, SVM, and BAL$^{-0.5}$ show similar classification results between the azimuthal angle of 180° and 270°. However, BAL$^{-0.5}$ misclassifies gates between 90° and 180° as convective type, which produces such high R^{CS}. On the other hand, MRMS and SVM show similar classification results in this region.

3.2.2 Tropical convective

Typhoon Morakot (6∼10 August 2009) brought significant rainfall to Taiwan. Over 700 people were reported dead in the storm, and the property loss was more than 3.3 billion USD. For most of the time during its landfall in Taiwan, the precipitation was classified as a mixture of tropical convective and tropical stratiform types. The performances of SVM, BAL^{0}, and BAL$^{-0.5}$ were validated using 96-hour data from 6 to 9 August 2009, where the results from 10 August 2009 were not included in the evaluation because no significant precipitation was observed from that day. The time series plots of R^{CS} (A), CSI (B), POD (C) and FAR (D) are shown in Figure 8. It could be found that scores of R^{CS}, CSI, and POD from the BAL based approaches is evidently lower than the results from SVM and MRMS, and the latter two show similar performance during these four days.

Classification results from BAL^{0}, BAL$^{-0.5}$, SVM (0402 UTC), and MRMS (0400 UTC) from 9 August 2009 are shown in Figure 9A, 9B, 9C, and 9D, respectively. The classification results ~~in those~~ within two regions, highlighted with two circles, are convective (SVM and MRMS) and stratiform (BAL^{0} and BAL$^{-0.5}$). Radar variables from 0402 UTC are shown in Figure 10 ~~includes~~ including the reflectivity (A), differential reflectivity (B), and separation index (C)~~from 0402 UTC, where~~ , respectively. Figure 10D is the zoom-in reflectivity field inside the red rectangular box (A) for more details. It was found that ~~the~~ a heavy precipitation band is on the top of RCMK (Figure 10D), and this may cause significant attenuation on Z and Z_{DR} fields. Although both Z and Z_{DR} fields were corrected using Equation 1, deficient or ~~over~~ excessive compensations on Z and Z_{DR} fields lead to increased uncertainty on the separation index. It may be the primary reason causing the small values of the separation index. Other reasons such as wet radome may also contribute to the Z and Z_{DR} issues. In Figure 10C, the separation index i are equal or less than -0.5 in the circled areas, and the BAL based approaches classify these regions as stratiform. On the other hand, these regions clearly show the convective precipitation features in the fields of Z (10A) and Z_{DR} (~~10D~~10B).

3.2.3 Stratiform precipitation event

The performances of BAL^0, $BAL^{-0.5}$, and SVM ~~approaches~~ were also evaluated with a widespread stratiform precipitation event on 26 March 2011. ~~There~~ This is a typical stratiform precipitation event, and there were no convective type ~~precipitations~~ pixels identified by MRMS~~, and all these~~ . These three approaches showed consistent classification results with the MRMS result during an 8-hour period evaluation. Since no convective type pixels were detected by these three approaches either, evaluation scores were not calculated and presented.

3.3 ~~Sensitivity test~~

~~The performances of BAL^0, $BAL^{-0.5}$ and proposed SVM were further validated respecting to the Z_{DR} bias. First, the impact of Z_{DR} bias on i is investigated through a simple simulation. In the simulation, the separation index i is calculated using Equations 2~5 with four distinct Z values: 10 dBZ, 20 dBZ, 30 dBZ, and 40 dBZ. For each Z value, Z_{DR} changes from -0.5 dB to 2 dB to simulate the Z_{DR} bias. The obtained i results are shown in Figure 11, and the symbol of triangle, diamond, cross, and pentagram indicates the result from 10 dBZ, 20 dBZ, 30 dBZ, and 40 dBZ, respectively. It could be found that for each Z, the calculated i drops when Z_{DR} increases. Moreover, a larger Z produces a larger i for the same Z_{DR} value. As introduced in Section 2.1, the precipitation may be classified as stratiform when i is less than 0. Therefore, positive Z_{DR} calibration bias may result in misclassifying more precipitation as stratiform type .~~

~~The impact of the Z_{DR} calibration bias on the performance of BAL^0, $BAL^{-0.5}$, and SVM was investigated using precipitation events from 30 August 2011. In this study, the Z_{DR} filed was first corrected from attenuation, and a ΔZ_{DR} was then manually added on the corrected Z_{DR} field as the artificial bias. The ΔZ_{DR} was set as: -0.2 dB, -0.1 dB, 0 dB, 0.1 dB, and 0.2 dB, respectively. The biased Z_{DR} was calculated as $Z_{DR}^b = Z_{DR} + \Delta Z_{DR}$. The separation index i was calculated using Z_{DR}^b through Equations 2~5, and classification results from BAL^0 and $BAL^{-0.5}$ were then calculated. The same trained weights and bias vector described in Section 2.3.2 were used in the SVM approach. Following the procedure described in Section 3.2, scores of R^{CS} (A), CSI (B), POD (C) and FAR (D) are calculated and shown in Figure 12. It should be noted that these scores are the 24-hour averaged values. It could be found that when the ΔZ_{DR} changes from -0.2 dB to 0.2 dB, the R^{CS} from both BAL^0 and $BAL^{-0.5}$ approaches decrease. This indicates that BAL^0 and $BAL^{-0.5}$ classify more precipitation as stratiform, and this results is consistent with the simulation. Both CSI and POD from BAL^0 and $BAL^{-0.5}$ show degradations with the increase of ΔZ_{DR}. On the other hand, the proposed SVM shows slightly better performances when ΔZ_{DR} changes from negative to positive. Both CSI and POD increase when ΔZ_{DR} increases, and the R^{CS} also has the similar trend. One possible reason is that convective type precipitation is normally associated with larger Z_{DR}. As a result, positive ΔZ_{DR} classify more precipitation as convective type. Similar results were also obtained from the case of 6~10 August 2009.~~

4 Conclusions

A novel precipitation classification approach using a support vector machine approach was developed and tested on a C-band polarimetric radar located in Taiwan. Different from other classification algorithms that use ~~whole~~ a complete volume scan data, the proposed method only utilizes the data from the lowest unblocked tilt to separate precipitation into convective or stratiform type. ~~It can be applied to~~ This feature makes this approach an optimal option in new scanning schemes ~~with more frequent scans at the lowest tilts and lack of information from a higher tilt,~~ such as AVSET, SAILS, MRLE, and etc. Three radar variables of reflectivity, differential reflectivity, and the separation index derived by Bringi et al. (2009) are utilized in the new proposed approach~~, where both~~. Both reflectivity and differential reflectivity need be corrected from attenuation and differential attenuation before applied in this approach. Although the separation index alone can be used in the precipitation classification, there ~~may be~~ are two potential limitations: thresholds and ~~the~~ biases on reflectivity and/or differential reflectivity. ~~Although the threshold~~ A threshold of "0" was suggested ~~to be used~~ in separating convective type from stratiform type, however, it was found that a single threshold may not be sufficient for all cases. Other thresholds (such as "-0.5" used in the current work), sometimes can produce better results than "0". ~~The biases may come from mis-calibration, attenuation, wet radome, blockage. Although~~ On the other hand, although both reflectivity and differential reflectivity should be corrected from attenuation before used in the separation index calculation, the correction biases on either filed may cause large uncertainty in the derived separation index and further lead to a wrong classification. ~~Other factors also may have impacts on the separation index.~~

This work attempts to propose a complementary method to enhance the performance of using the separation index only. The proposed approach integrates input variables with a support vector machine method. The ~~weighs and bias vectors~~ weights vector and bias used in the support vector machine were trained with typical stratiform and convective precipitation events. It should be noted that the proposed approach has a flexible framework, and some other variables can be easily included. With ~~newly~~ new added variables, the ~~weighting and bias vectors~~ weight vector and bias need to be retrained. ~~The~~ In the current work, the proposed approach was tested with multiple ~~cases~~precipitation events. Its performance was found better than using then separation index only and similar to a ~~well-developed~~ well developed approach, MRMS, which utilizes multiple tilts radar data in the classification. It should be noted that ~~the time difference between RCMK (i.e., BAL^0, BAL$^{-0.5}$, and SVM) and MRMS could be as large as 5 minutes. Therefore, the pixel-to-pixel evaluation criteria of the critical success index (CSI), probability of detection (POD) and false alarm rate (FAR) may not really reflect their performances. Although a new variable of~~ although the proposed approach shows better scores (POD, FAR, CSI and R^{CS}~~is used in the performance evaluation, this~~), this evaluation should be treated as qualitative evaluation ~~.~~ instead of statistical analysis. In order to obtained statistical evaluation results, more long-term precipitation events are needed.

There are some issues that need to be noticed before applying this approach into operation. First, this approach is developed for fast ~~scanning~~ scan and fast update purpose, ~~therefore,~~ and data from the lowest unblocked tilt is used as the input. ~~However, if~~ If the radar is located in a complex orography area, radar ~~beam~~ beams could be partially or completely blocked at some regions. A possible solution for such scenario is using data from different ~~scanning~~ tilts to form a hybrid scan, and the hybrid scan is then used as the input. ~~Radar scanning tilts used in the hybrid scanning are determined by the radar scanning geometry.~~

~~Given the factor~~ Given the fact that precipitation's microphysics (such as drop size distribution) from different altitudes may be significantly different, ~~therefore,~~ the performance of proposed approach may be worse than expected. Second, the performance of the proposed approach depends highly on the training data, which should be selected very careful. In the current work, a threshold of $\rho_{HV} > 0.9$ is used in the data selection. Using lower threshold may cause different performance, and more investigations on this issue are needed. Third, coefficients in the separation index calculation depend on the local drop size distribution and drop shape relation features. Therefore, new relations need to be derived for the optimal results. Moreover, the separation index only validates at liquid phase precipitation. For ice phase precipitation, such as mixed hail and rain, its performance is not well studied. Other hydrometeor classification schemes could be used for such scenario. Fourth, the mis-calibration may significantly affect the performance of the proposed approach. The calibration biases for Z and Z_{DR} should be within 1 dBZ and 0.2 dB, respectively. Moreover, this work only presents a prototype algorithm. Given the flexible framework, other variables (such as differential phase) could be easily integrated into this algorithm, and the performance could be further enhanced.

Code and data availability.

The datasets and source code used in this study are available from the corresponding author upon request (yadwang@siue.edu).

Author contributions.

The algorithm was originally developed by Dr. Y. Wang. Dr. L. Tang processed the radar data including generate results from MRMS. Dr. P.-L. Chang and Miss Y.-S. Tang provided and processed radar data from CWB, they were further involved in algorithm discussion and article writing.

Competing interests.

The authors declare that there is no conflict of interest.

Acknowledgements. The authors thank the radar engineers from CWB help us collecting and processing the radar data.

References

Adler, R. F. and Negri, A. J.: A satellite infrared technique to estimate tropical convective and stratiform rainfall, J. Appl. Meteor., 27, 30–51, 1988.

Anagnostou, E. N.: Doppler radar characteristics of precipitation at vertical incidence, Rev. Geophys. Space Phys., 11, 1–35, 2004.

Biggerstaff, M. I. and Listemaa, S. A.: An improved scheme for convective/stratiform echo classification using radar reflectivity, J. Appl. Meteor., 39, 2129–2150, 2000.

Bringi, V. N., Chandrasekar, V., Hubbert, J., Gorgucci, E., Randeu, W. L., and Schoenhuber, M.: Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis, J. Atmos. Sci., 60, 354–365–2122, 2003.

Bringi, V. N., Williams, C. R., Thurai, M., and May, P. T.: Using dual-polarized radar and dual-frequency profiler for DSD characterization: a case study from Darwin, Australial, J. Atmos. Oceanic Technol., 26, 2107–2122, 2009.

Burges, C. J. C.: A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discovery, 2, 955–974, 1998.

Capozzi, V., Montopoli, M., Mazzarella, V., Marra, A. C., Panegrossi, N. R. G., Dietrich, S., and Budillon, G.: Multi-variable classification approach for the detection of lightning activity using a low-cost and portable X band radar, Remote Sens., 10, 1797, 2018.

Carey, L. D., Rutledge, S. A., Ahijevych, D. A., and Keenan, T. D.: Correcting propagation effects in C-band polarimetric radar observations of tropical convection using differential propagation phase, J. Appl. Meteor., 39, 1405–1433, 2000.

Chang, W.-Y., Wang, T.-C. C., and Lin, P.-L.: Characteristics of the raindrop size distribution and drop shape relation in typhoon systems in the western Pacific from the 2D video disdrometer and NCU C-band polarimetric radar., J. Atmos. Oceanic Technol., 26, 1973–1993, 2009.

Gourley, J. J., Flaming, Z. L., Vergara, H., Kirstetter, P.-E., Clark, R. A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J. M., Hong, Y., and Howard, K.: The FLASH project: improving the tools for flash flood monitoring and prediction across the United States, Bull. Amer. Meteor. Soc., 94, 799–805, 2016.

Grecu, M., Olson, W. S., Munchak, S. J., Ringerud, S., Liao, L., Haddad, Z. S., Kelley, B. L., and McLaughlin, S. F.: The GPM combined algorithm, J. Atmos. Oceanic Technol., 33, 2225–2245, 2016.

Haykin, S. O., ed.: Neural networks and learning machines, Pearson Higher Ed., PP 936, 2011.

Hong, Y., Kummerov, C. D., and Olson, W. S.: Separation of convective and stratiform precipitation using microwave brightness temperature, J. Appl. Meteor., 38, 1195–1213, 1999.

Houghton, H. G.: On precipitation mechanisms and their artificial modification, J. Appl. Meteor., 7, 851–859, 1968.

Houze, R. A. J., ed.: Cloud Dynamics, Academic Press, PP. 573, San Diego, 1993.

Houze, R. L.: Stratiform precipitation in regions of convection: A meteorological paradox?, Bull. Amer. Meteor. Soc., 78, 2179–2196, 1997.

Jameson, A. R.: The effect of temperature on attenuation correction schemes in rain using polarization propagation differential phase shift, J. Appl. Meteor., 31, 1106–1118, 1992.

Kumjian, M. R.: Principles and applications of dual-polarization weather radar. Part I description of the polarimetric radar variables., J. Operational. Meteor., 19, 226–242, 2013.

Leary, C. A. and Jr., R. A. H.: Melting and evaporation of hydrometeors in precipitation from the anvil clouds of deep tropical convection, J. Atmos. Sci., 36, 669–679, 1979.

Park, H., Ryzhkov, A. V., Zrnić, D. S., and Kim, K.-E.: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS, Wea. Forecasting, 24, 730–748, 2009.

Park, S. G., Maki, M., Iwanami, K., Bringi, V. N., and Chandrasekar, V.: Correction of radar reflectivity and differential reflectivity for rain attenuation at X-band. part II: evaluation and application, J. Atmos. Oceanic Technol., 22, 1633–1655, 2005.

Penide, G., Protat, A., Kumar, V. V., and May, P. T.: Comparison of two convective/stratiform precipitation classification techniques: radar reflectivity texture versus drop size distribution-based approach, J. Atmos. Oceanic Technol., 30, 2788–2797, 2013.

Powell, S. W., Jr., R. A. H., and Brodzik, S. R.: Rainfall-type categorization of radar echoes using polar coordinate reflectivity data, J. Atmos. Oceanic Technol., 33, 523–538, 2016.

Ryzhkov, A. V. and Zrnic, D. S., eds.: Radar Polarimetry For Weather Observations, Springer Atmospheric Sciences, PP 477, 2019.

Skofronick-Jackson, G. and Coauthors: The global precipitation measurement GPM mission for science and society, Bull. Amer. Meteor. Soc., 98, 1679–1695, 2017.

Smith, T. M., Lakshmanan, V., Stumpf, G. J., Ortega, K. L., Hondl, K., Cooper, K., Calhoun, K. M., Kingfield, D. M., Manross, K. L., Toomey, R., and Brogden, J.: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities, Bull. Amer. Meteor. Soc., 97, 1617–1630, 2016.

Steiner, M., Jr., R. A. H., and Yuter, S. E.: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data, J. Appl. Meteor., 34, 1978–2007, 1995.

T., A., Somula, R., K., G., Saxena, A., and A., P.: Estimating rainfall using machine learning strategies based on weather radar data, Int J Commun Syst, p. e3999, https://doi.org/10.1002/dac.3999, 2019.

Testud, J., Bouar, E. L., Obligis, E., and Ali-Mehenni, M.: The rain profiling algorithm applied to polarimetric weather radar, J. Atmos. Oceanic Technol., 17, 332–356, 2000.

Thompson, E. J., Rutledge, S. A., Dolan, B., and Thursai, M.: Drop size distributions and radar observations of convective and stratiform over the equatorial Indian and West Pacific Oceans, J. Atmos. Sci., 72, 4091–4125, 2015.

Tokay, A. and Short, D. A.: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds, J. Appl. Meteor., 35, 355–371–4125, 1996.

Vulpiani, G., Montopoli, M., Passeri, L. D., Gioia, A. G., Giordano, P., and marzano, F. S.: On the use of dual-polarized C-band radar for operational rainfall retrieval in mountainous areas, J. Appl. Meteor. Climatol., 51, 405–425, 2012.

Wang, Y., Zhang, J., Ryzhkov, A. V., and Tang, L.: C-band polarimetric radar QPE based on specific differential propagation phase for extreme typhoon rainfall, J. Atmos. Oceanic Technol., 30, 1354–1370, 2013.

Wang, Y., Zhang, P., Ryzhkov, A. V., Zhang, J., and Chang, P.-L.: Utilization of specific attenuation for tropical rainfall estimation in complex terrain, J. of Hydrometeorology, 15, 2250–2266, 2014.

Yang, Y., Chen, X., and Qi, Y.: Classification of convective/stratiform echoes in radar reflectivity observations using a fuzzy logic algorithm, J. Geophys. Res. Atmos., 118, 1896–1905, 2013.

Yen, M., Liu, D., and Hsin, Y.: Application of the deep learning for the prediction of rainfall in Southern Taiwan, Sci. Rep., 9, 12 774, 2019.

Zhang, G., ed.: Weather Radar Polarimetry, CRC Press PP 304, 2016.

Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., Cooten, S. V., Kitzmiller, K. K. D., Ding, F., Seo, D.-J., Wells, E., and Dempsey, C.: National mosaic and multi-sensor QPE (NMQ) system: Description, results, and future plans, Bull. Amer. Meteor. Soc., 92, 1321–1338, 2011.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., and Kitzmiller, D.: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: initial operating capabilities, Bull. Amer. Meteor. Soc., 97, 621–638, 2016.

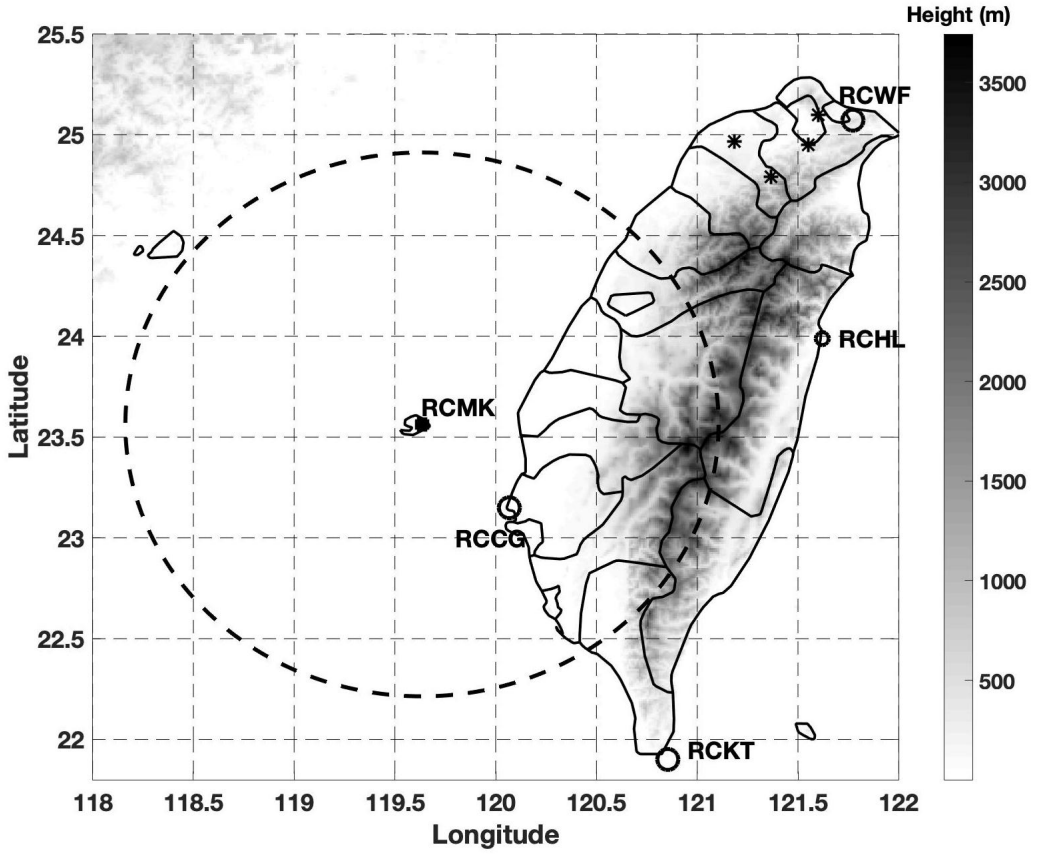

Figure 1. The terrain of Taiwan, the location of a C-band polarimetric radar RCMK (marked with a black square), JWDs (marked with black stars), and four S-band single-polarization radar RCCG, RCKT, RCHL, and RCWF (marked with black circles). The continuous grey-scale terrain map shows the central mountain range of Taiwan.

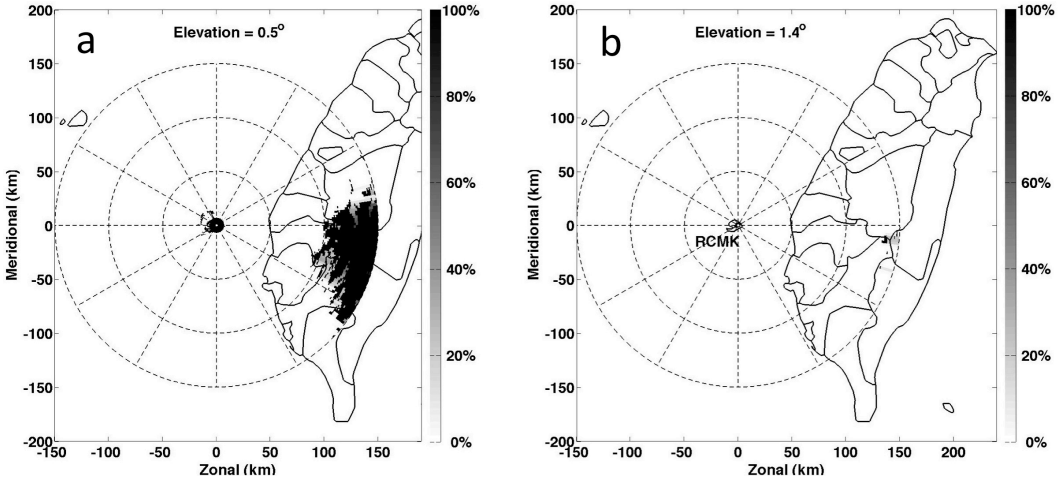

Figure 2. Blockage maps of RCMK from the first 2 EAs (0.5° and 1.4°). The grey scale indicates the blockage percentages.

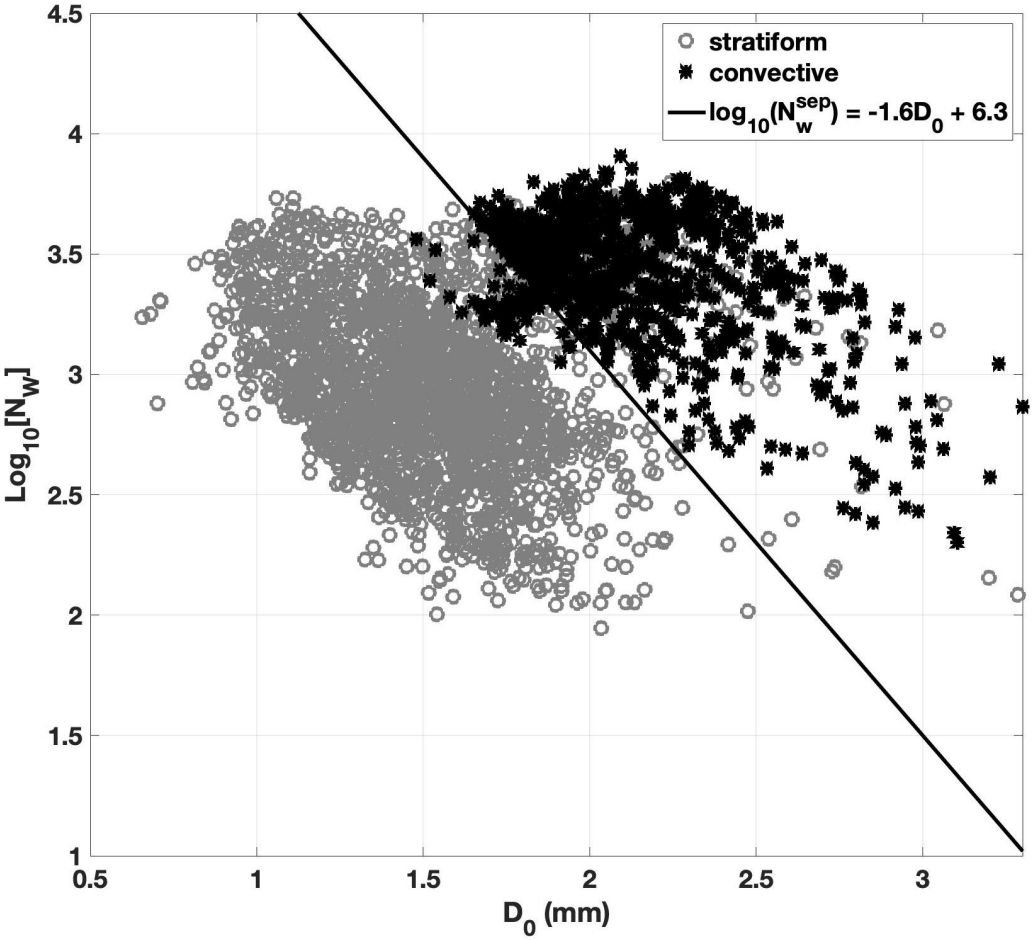

Figure 3. The distribution of $log10(N_w)$ vs D_0. The DSD data from stratiform and convective precipitations are presented with gray circles and black stars, and the separator line is shown with a solid line.

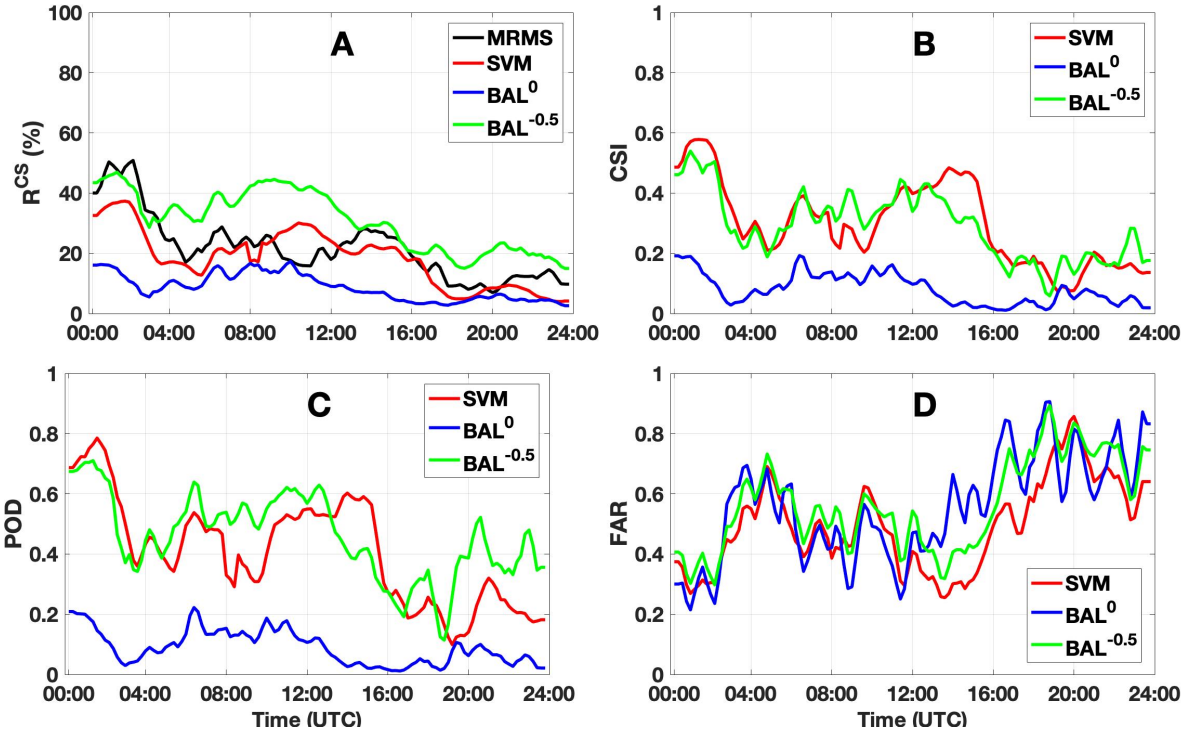

Figure 4. The time series plot of R^{CS}(A), CSI(B), POD(C), and FAR(D) from 30 August 2011. 24-hours data 0000 UTC 2400 UTC are used in each case. The results from BAL with threshold T_0 = -0.5, BAL with threshold T_0 = 0, SVM, MRMS, are presented by green, blue, red and black lines, respectively.

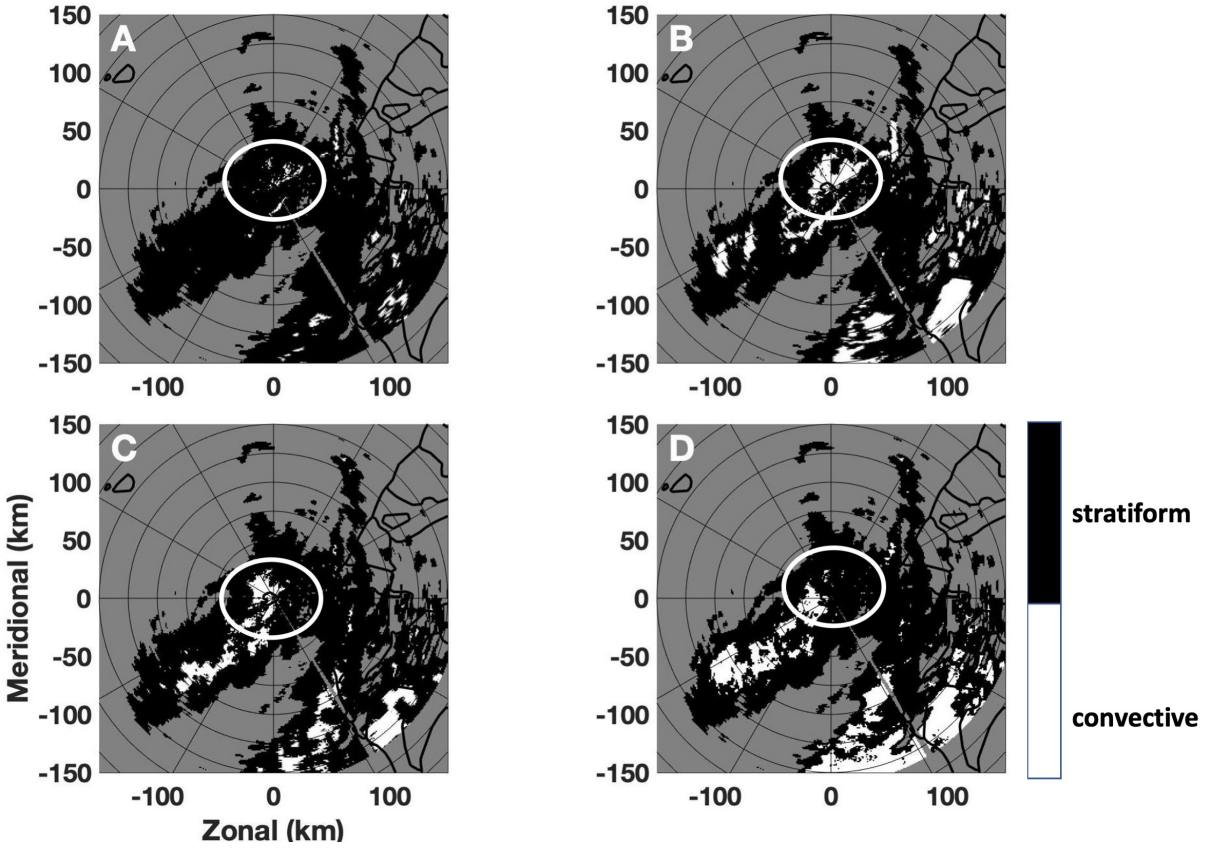

Figure 5. The classification results from BAL^0(A), $BAL^{-0.5}$(B), SVM(C) and MRMS(D). The time stamp for BAL^0, $BAL^{-0.5}$, and SVM is 0303 UTC 30 August 2011, and time stamp for MRMS is 0300 UTC 30 August 2011. The region inside the white circle is used in the analysis.

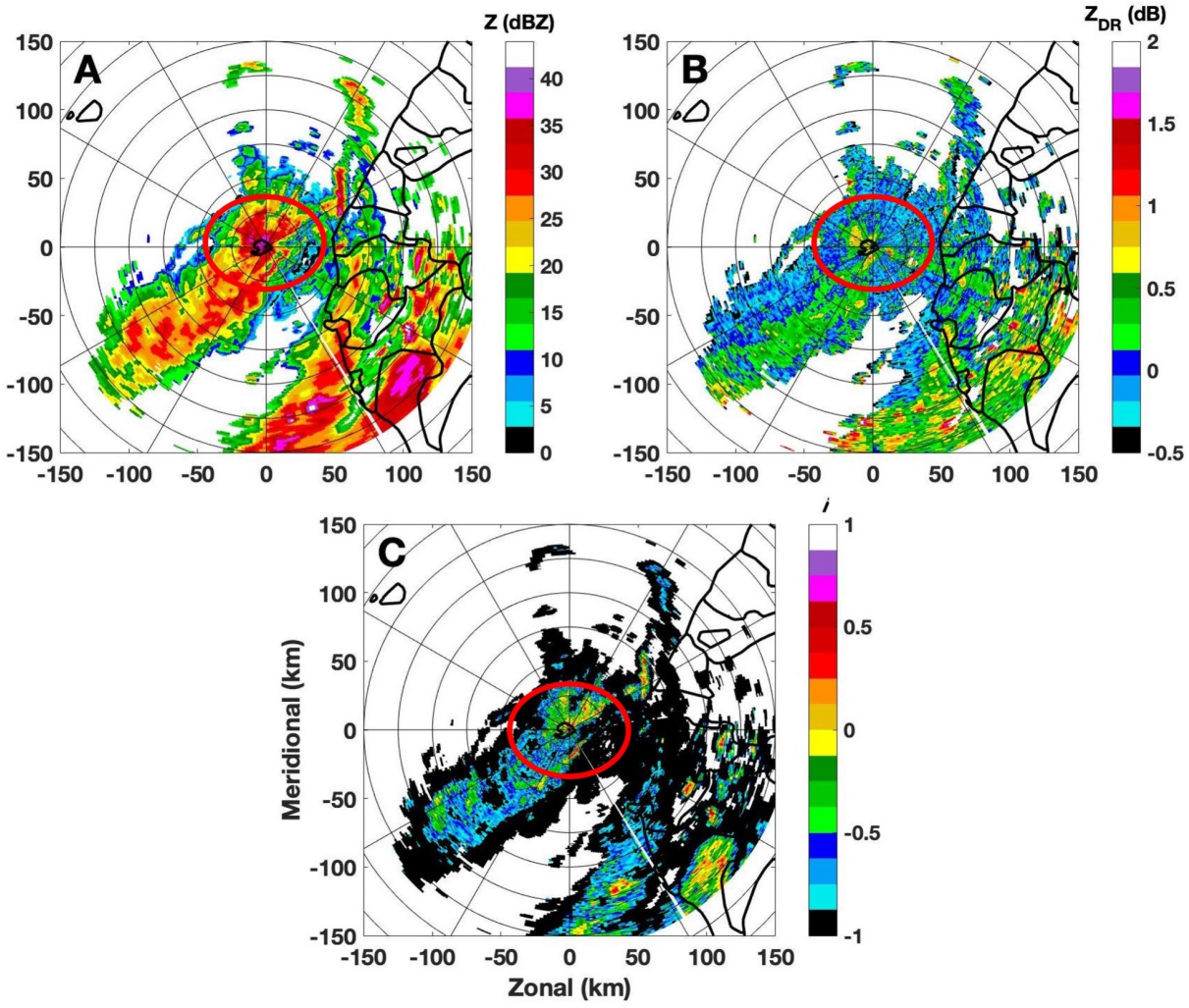

Figure 6. Radar variables of reflectivity (A), differential reflectivity(B), and separation index(C). The radar data was collected by RCMK at 0303 UTC 30 August 2011. The region inside the red circle is used in the analysis.

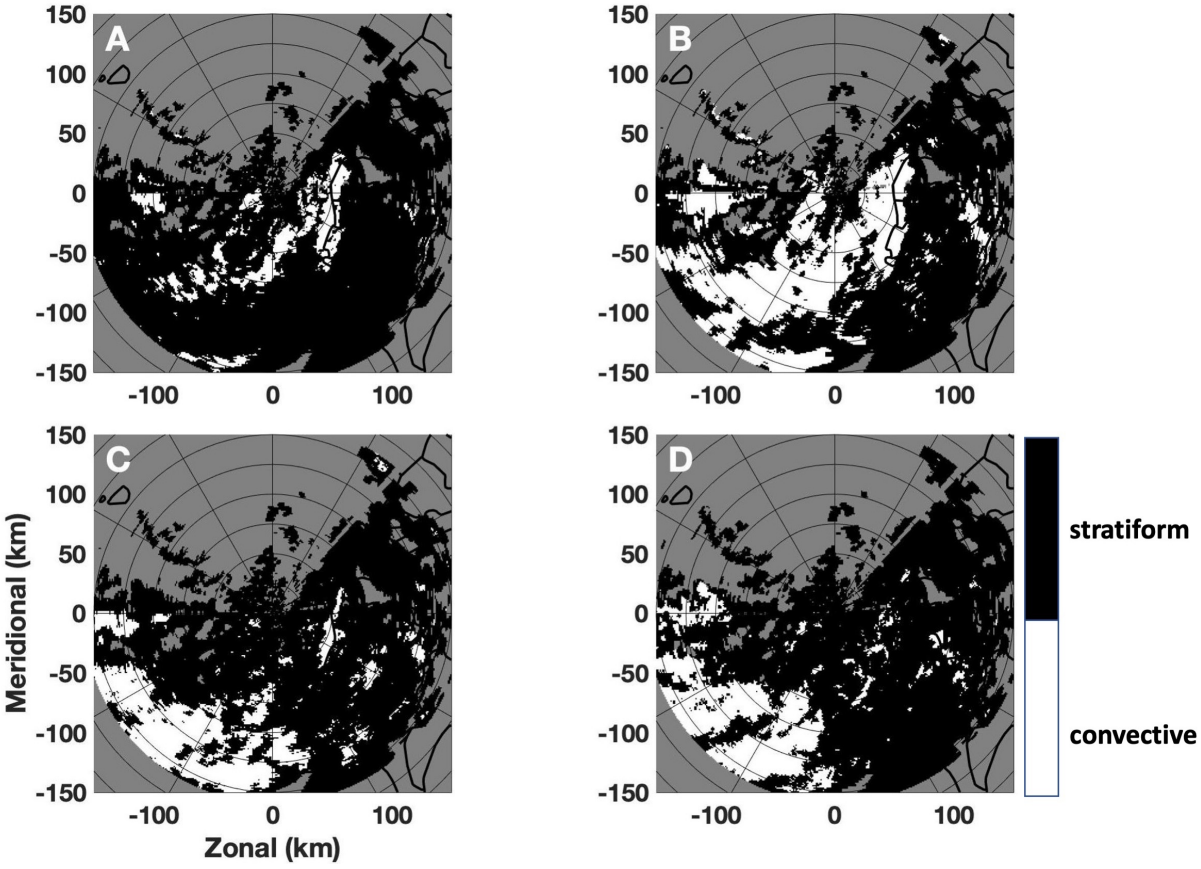

Figure 7. Similar to Figure 5, results are from 0650 UTC.

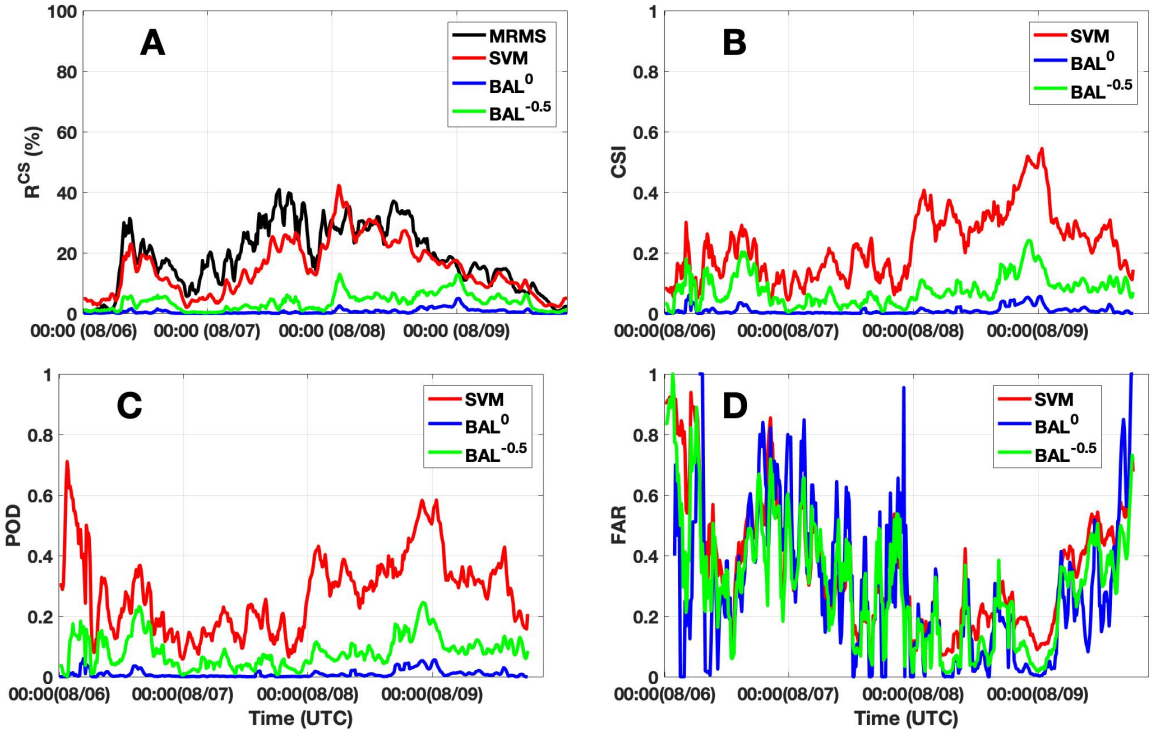

Figure 8. The time series plot of R^{CS}(A), CSI(B), POD(C), and FAR(D) from 06∼09 August 2009. 96-hours data are used in each case. The results from BAL with threshold T_0 = -0.5, BAL with threshold T_0 = 0, SVM, and MRMS are indicated by green, blue, red and black lines, respectively.

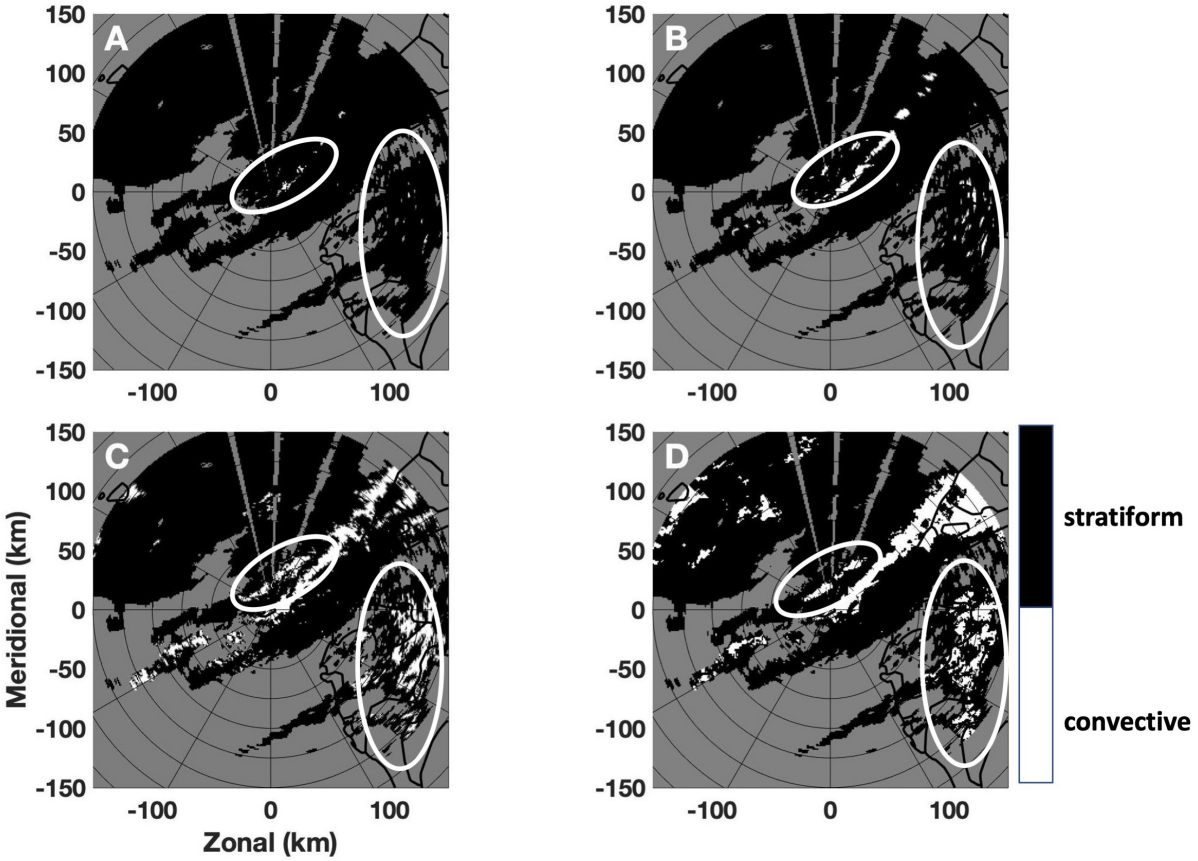

Figure 9. The classification results from BAL^{0}(A), $BAL^{-0.5}$(B), SVM(C), and MRMS(D). The time stamp for BAL^{0}, $BAL^{-0.5}$, and SVM is 0402 UTC 9 August 2009, and time stamp for MRMS is 0400 UTC 9 August 2009.

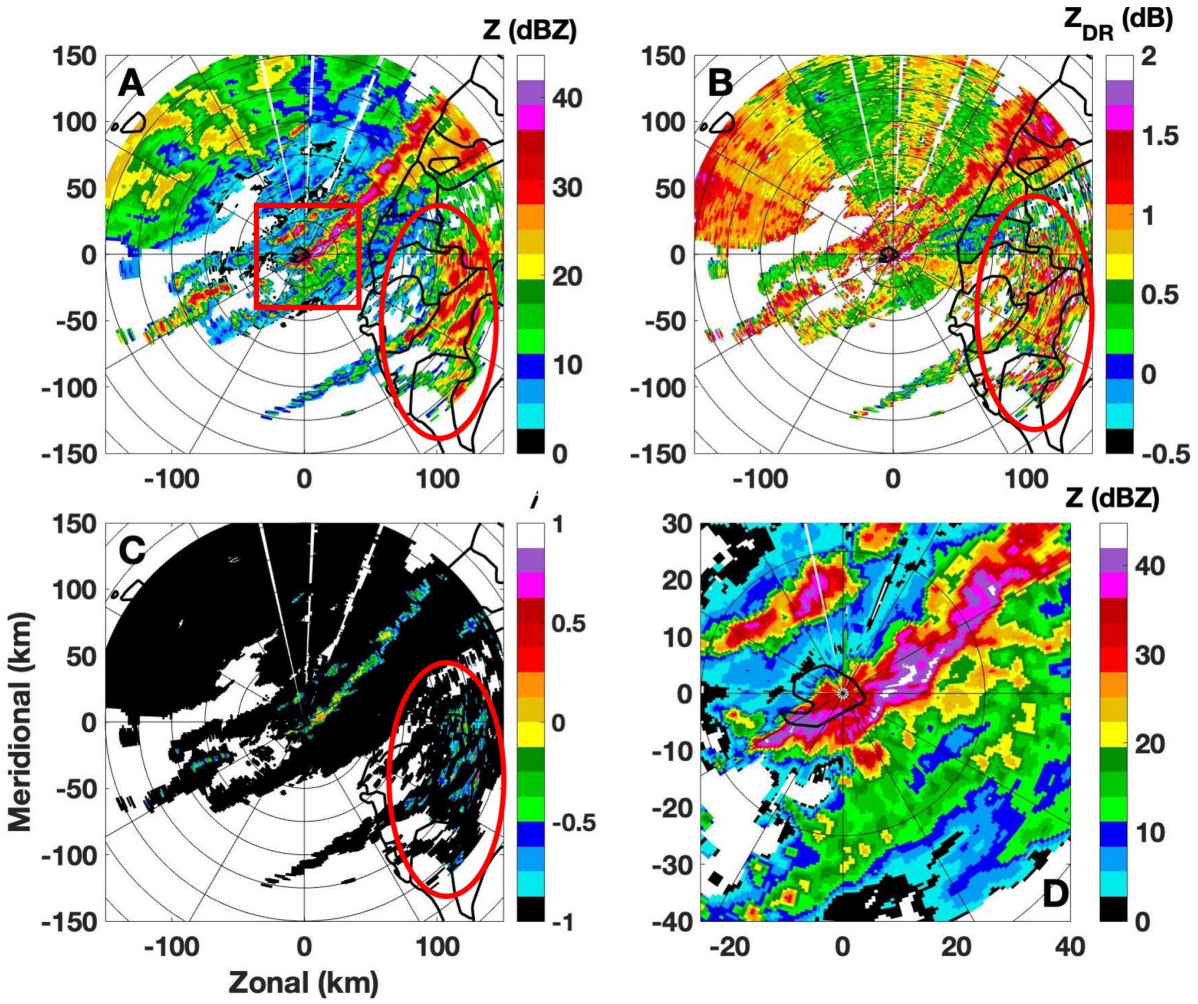

Figure 10. Radar variables of reflectivity(A), differential reflectivity(B), separation index(C), and reflectivity within the red rectangular box in A(D). The radar data was collected by RCMK at 0402 UTC 9 August 2009.

~~The calculated separation index respecting to different differential reflectivity values.~~

~~24-hour averaged R^{CS}(A), CSI(B), POD(C), and FAR(D) from 30 August 2011. The results from BAL with threshold T_0 = -0.5, BAL with threshold T_0 = 0, SVM, and MRMS are indicated with symbols of pentagram, circle, triangle, and square, respectively.~~